(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 April 2005 (14.04.2005)

PCT

(10) International Publication Number
WO 2005/032328 A2

(51) International Patent Classification[7]: A61B

(21) International Application Number: PCT/US2004/015761

(22) International Filing Date: 20 May 2004 (20.05.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/472,330 21 May 2003 (21.05.2003) US

(71) Applicant *(for all designated States except US)*: **MILLENNIUM PHARMACEUTICALS, INC.** [US/US]; 40 Landsdowne Street, Cambridge, MA 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **GUILD, Braydon, C.** [US/US]; 109 Riverdale Road, Concord, MA 01742 (US). **LIAO, Hua** [CN/US]; 15 Damon Park, Apt.1, Concord, MA 02474 (US). **JONES, Michael, D.** [US/US]; 14 Schouler Court, Arlington, MA 02476 (US). **WU, Jiang** [CN/US]; 23 Judith Lane, Apartment 6, Waltham, MA 02452 (US). **ZOLG, Johannes, W.** [DE/DE]; Sportplatzweg 4A, D-82372 Weilheim (DE).

(74) Agents: **SILVERI, Jean, M.** et al.; Millennium Pharmaceuticals, Inc., 40 Landsdowne Street, Cambridge, MA 02139 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF RHEUMATOID ARTHRITIS

(57) **Abstract:** The invention relates to compositions, kits, and methods for detecting, characterizing, preventing, and treating human Rheumatoid Arthritis (RA). A variety of newly-identified markers are provided, wherein changes in the levels of expression of one or more of the markers is correlated with RA.

WO 2005/032328 A2

COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF RHEUMATOID ARTHRITIS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application Number 60/472,330, filed May 21, 2003, the contents of which are incorporated herein by this reference.

FIELD OF THE INVENTION

[0002] The field of the invention is rheumatoid arthritis, including diagnosis, prognosis, characterization, management, and therapy of rheumatoid arthritis.

BACKGROUND OF THE INVENTION

[0003] Rheumatoid arthritis ("RA") is a chronic, inflammatory, systemic disease that produces its most prominent manifestations in the diarthrodial joints. Persistent and progressive synovitis develops in peripheral joints. RA encompasses a wide spectrum of features, from self-limiting disease to progressively chronic disease with varying degrees of joint destruction to clinically evident extra-articular manifestations. Genetic and environmental factors control the progression, extent, and pattern of the inflammatory response and are thereby responsible for the heterogeneous clinical features.

[0004] RA has a worldwide distribution and involves all ethnic groups. Although the disease can occur at any age, the prevalence increases with age and the peak incidence is between the fourth and sixth decade, although data from population- based prevalence and incidence studies have to be interpreted cautiously because there is no laboratory test, histologic finding, or radiographic-feature to confirm a diagnosis of RA.

[0005] The most widely used system to classify RA is the American College of Rheumatology 1987 revised criteria for the classification of RA. Arnett FC, *et al.*, 1988, The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 31:315-324. According to the criteria, a patient is said to have RA if the patient satisfies at least four of the following seven criteria and criteria 1-4 must be present for at least six weeks: 1) morning stiffness, 2) arthritis of three or more joint areas, 3) arthritis of hand joints, 4) symmetrical arthritis, 5) rheumatoid nodules, 6) serum rheumatoid factor ("RF"), and 7) radiographic changes. These criteria

have a sensitivity and specificity of approximately 90%. Depending on the stringency of the criteria, prevalence vary from 0.3% to 1.5% in the North American population. The prevalence is about 2.5 times higher in females than in males.

[0006] The histologic changes in RA are not disease-specific but largely depend on the organ involved. The primary inflammatory joint lesion involves the synovium. The earliest changes are injury to the synovial microvasculature with occlusion of the lumen, swelling of endothelial cells, and gaps between endothelial cells, as documented by electron microscopy. This stage is usually associated with mild proliferation of the superficial lining cell layer. Two cell types constitute the synovial lining: bone marrow-derived type A synoviocyte, which has macrophage features, and mesenchymal type B synoviocyte. Both cell types contribute to the synovial hyperplasia, suggesting a paracrine interaction between these two cell types. This stage of inflammation is associated with congestion, edema, and fibrin exudation. Cellular infiltration occurs in early disease and initially consists mainly of T lymphocytes. As a consequence of inflammation, the synovium becomes hypertrophic from the proliferation of blood vessels and synovial fibroblasts and from multiplication and enlargement of the synovial lining layers. Granulation tissue extends to the cartilage and is known as pannus. The tissue actively invades and destroys the periarticular bone and cartilage at the margin between synovium and bone, known as erosive RA.

[0007] The articular manifestations of RA can be placed in two categories: reversible signs and symptoms related to inflammatory synovitis and irreversible structural damage caused by synovitis. This concept is useful not only for staging disease and determining prognosis but also for selecting medical or surgical treatment. Structural damage in the typical patient usually begins sometime between the first and second year of the disease. Van der Heijde, DM, *et al.*, 1982, *Arthritis Rheum* 25:361-365. Although synovitis tends to follow a fluctuating pattern, structural damage progresses as a linear function of the amount of prior synovitis.

[0008] The etiology of the early events in RA remains elusive. The possibility of a bacterial or viral infection has been vigorously pursued. All efforts to associate an infectious agent with RA by isolation, electron microscopy, or molecular biology have failed. It is possible that there is no single primary cause of RA and that different mechanisms may lead to the initial tissue injury and precipitate synovial inflammation.

[0009] Clinical signs of synovitis may be subtle and are often subjective. Warm, swollen, obviously inflamed joints are usually seen only in the most active phases of

inflammatory synovitis. Cartilage loss and erosion of periarticular bone are the characteristic features of structural damage. The clinical features related to structural damage are marked by progressive deterioration functionally and anatomically. Structural damage to the joint is irreversible and additive.

[0010] Data from longitudinal clinical and epidemiologic studies provide guidelines for treatment. These studies emphasize 1) the need for early diagnosis, 2) identification of prognostic factors, and 3) early aggressive treatment. Earlier diagnosis and treatment, preferably within the first several months after onset of symptoms, may help prevent irreversible joint damage. The present invention provides such methods and reagents for the diagnosis, characterization, prognosis, monitoring and treatment of RA.

SUMMARY OF THE INVENTION

[0011] The present invention is directed to the methods of determining or diagnosing whether patients are afflicted with inflammatory disorders, e.g., joint disorders, i.e., rheumatoid arthritis (RA). These methods typically include the step of obtaining a sample of a patient's bodily fluid, e.g., blood serum, determining the level of expression of one or more markers in the fluid, and identifying whether the patient's body fluid has a pattern or profile or expression of a selected marker or marker set (a pattern or profile of expression is also referred to herein as the "expression" or "marker profile" of the marker set.) which correlates with the presence of an inflammatory disorder.

[0012] The present invention also provides methods for determining or diagnosing whether patients are afflicted with a particular form of arthritis, i.e., erosive RA. Erosive RA is characterized by erosions or pits in the surface of the bone adjacent to the articular surface. In particular, in erosive RA, the granulation tissue actively invades and destroys the periarticular bone and cartilage at the margin between the synovium and the bone. These methods typically include the step of obtaining a sample of a patient's bodily fluid, e.g., blood serum, determining the level of expression of one or more markers in the fluid, and identifying whether the patient's body fluid has a pattern or profile or expression of a selected marker or marker set which correlates with the presence of erosive or non-erosive RA. The present invention therefore provides methods, reagents and kits for diagnosing, characterizing, prognosing, monitoring, and treating RA, including identifying erosive and non-erosive RA.

[0013] In the methods of the present invention, the samples or patient samples may comprise RA-associated body fluids. Such fluids include, for example, blood fluids, (*e.g.*,

whole blood, blood serum, plasma, blood having platelets removed there from, etc.), urine, saliva, tears, and synovial fluid. The patient samples may also comprise cells, *e.g.*, cells obtained from the patient. The cells may be endothelial cells, white blood cells and synovium cells, osteoclasts, osteoblasts, chondrocytes as well other cells found in joints. In a further embodiment, the patient sample is *in vivo*.

[0014] The markers of the invention, whose expression correlates with the presence or absence of RA, are identified in Table 1 (herein after identified as "RA markers" or "markers"). The markers in Table 1 were identified by the sequencing of peptides derived from proteins in the sera of healthy, non-erosive and erosive patients by mass spectroscopy (see Experimental Protocol section below). Table 1 headings used are marker identification number ("Marker #"), the name the marker is commonly known by, if applicable ("Gene Name"), the data generated from each serum sample ("Erosive", "Non-Erosive", and "Healthy"), the corresponding molecular weight or the marker ("Protein MW (Da)"), the corresponding GenBank GI Number of the marker ("accession number"). Table 1 lists data collected for each marker. The heading "# of spectra" is defined as the number of peptides detected from a particular marker. The heading "total intensity" is defined as a measure used for marker quantitation calculated as the sum of parent m/z abundance in the MS scans, (~ chromatographic peak area), dependent upon the user designated scan tolerance (scan number separation in chromatographic time), the putative parent m/z (as adjusted by user designation of Find parent ^{12}C) and the user designated mass tolerance allowed for merging scans with the same parent m/z. The total intensity is summed so that each observation of a peptide counts towards the total intensity for the marker. Once data was collected from all three pools of patient sera, the data output was aligned and visually inspected. Candidate markers were selected based upon assessing which proteins showed the largest number of spectra as well as the total intensity in RA samples (erosive and non-erosive) versus healthy samples.

[0015] Candidate markers are listed in Table 2. Table 2 headings used are marker identification number ("Marker #"), the name the marker is commonly known by, if applicable ("Gene Name"), the data generated from each serum sample ("Erosive", "Non-Erosive", and "Healthy"), the ratio of total intensities of a marker in erosive and non-erosive serum ("E:N"), the ratio of total intensities of a marker in erosive and healthy serum ("E:H"), the corresponding molecular weight or the marker ("Protein MW (Da)"), the corresponding GenBank GI Number of the marker ("accession number"), and where indicated, the sequence listing identifier of the cDNA sequence of a nucleotide transcript

encoded by or corresponding to the marker ("SEQ ID NO (nts)") and the sequence listing identifier of the amino acid sequence of a protein encoded by or corresponding to the marker ("SEQ ID NO (AA)"). By examining the expression of one or more of the identified markers or marker sets in a patient's serum, it is possible to determine whether a patient has RA or has higher than normal risk for developing RA. Also, by examining the expression of one or more of the identified markers or marker sets in a patient's serum, it is possible to determine whether a patient has erosive RA or has higher than normal risk for developing erosive RA.

[0016] According to the invention, any of the aforementioned methods may be performed using a plurality (*e.g.* 2, 3, 5, or 10 or more) of RA markers, including RA markers known in the art. In such methods, the level of expression in the sample of each of a plurality of markers, at least one of which is a marker of the invention, is compared with the normal level of expression of each of the plurality of markers in samples of the same type obtained from control humans not afflicted with RA. A significantly altered (*i.e.*, increased or decreased as specified in the above-described methods using a single marker) level of expression in the sample of one or more markers of the invention, or some combination thereof, relative to that marker's corresponding normal or control level, is an indication that the patient is afflicted with RA. For all of the aforementioned methods, the marker(s) are preferably selected such that the positive predictive value of the method is at least about 10%.

[0017] According to the invention, the marker(s) are selected such that the positive predictive value of the methods of the invention is at least about 10%, preferably about 25%, more preferably about 50% and most preferably about 90%. Also preferred are embodiments of the method wherein the marker is over- or under-expressed by at least two-fold in at least about 20% of fast-progressing RA.

[0018] In accordance with the methods of the present invention, the level of expression of the marker in a sample can be assessed, for example, by detecting the presence in the sample of:

a marker protein (*e.g.*, a protein having a sequence selected from the group consisting of the markers listed in Tables 1 and 2), or a fragment of the protein (*e.g.* using a reagent, such as an antibody, an antibody derivative, or an antibody fragment, which binds specifically with the marker protein or a fragment of the protein)

a metabolite which is produced directly (*i.e.*, catalyzed) or indirectly by a marker protein

a transcribed polynucleotide (*e.g.* an mRNA or a cDNA, including a polynucleotide selected from the group consisting of the markers listed in Tables 1 and 2) or fragment thereof, having at least a portion with which the marker nucleic acid is substantially homologous (*e.g.* by contacting a mixture of transcribed polynucleotides obtained from the sample with a substrate having one or more of the marker nucleic acids fixed thereto at selected positions)

a transcribed polynucleotide or fragment thereof, wherein the polynucleotide anneals with the marker nucleic acid under stringent hybridization conditions.

[0019] In a further aspect, the invention provides an antibody, an antibody derivative, or an antibody fragment, which binds specifically with a marker protein or a fragment of the protein. The invention also provides methods for making such antibody, antibody derivative, and antibody fragment. Such methods may comprise immunizing a mammal with a protein or peptide comprising the entirety, or a segment of 10 or more amino acids, of a marker protein, wherein the protein or peptide may be obtained from a cell or by chemical synthesis. The methods of the invention also encompass producing monoclonal and single-chain antibodies, which would further comprise isolating splenocytes from the immunized mammal, fusing the isolated splenocytes with an immortalized cell line to form hybridomas, and screening individual hybridomas for those that produce an antibody that binds specifically with a marker protein or a fragment of the protein.

[0020] In one aspect, the invention relates to various diagnostic, monitoring, test and other methods related to RA detection and therapy. In one embodiment, the invention provides a diagnostic method of assessing whether a patient has RA or has higher than normal risk for developing RA, comprising the steps of comparing the level of expression of a marker of the invention in a patient sample and the normal level of expression of the marker in a control, *e.g.*, a sample from a patient without RA or the expression level of the marker in a population-average. A significantly higher level of expression of the marker in the patient sample as compared to the normal level is an indication that the patient is afflicted with RA or has higher than normal risk for developing RA. It will be appreciated that the "level of expression" includes a quantitative measurement, *i.e.*, the sample may be analyzed quantitatively, wherein the abundance of one or more of the markers in a sample is determined and compared to the normal abundance of the one or more markers.

[0021] The methods of the present invention are particularly useful for patients with identified inflammatory synovitis or other symptoms associated with RA. The methods of the present invention can also be of particular use with patients having an enhanced risk of developing RA (*e.g.*, patients having a familial history of RA, patients identified as having a RF, patients at least about 40-60 years of age and female patients at least about 40-60 years of age). The methods of the present invention may further be of particular use in monitoring the efficacy of treatment of a RA patient (*e.g.* the efficacy of nonsteroidal anti-inflammatory drugs (NSAIDs), corticosteroids, and disease-modifying antirheumatic drugs (DMARDs)).

[0022] In another aspect, the invention relates to various diagnostic and test kits. In one embodiment, the invention provides a kit for assessing whether a patient is afflicted with RA. The kit comprises a reagent for assessing expression of a marker of the invention. In another embodiment, the invention provides a kit for assessing the suitability of a chemical or biologic agent for inhibiting RA in a patient. Such a kit comprises a reagent for assessing expression of a marker of the invention, and may also comprise one or more of such agents. Such kits may comprise an antibody, an antibody derivative, or an antibody fragment, which binds specifically with a marker protein, or a fragment of the protein. Such kits may also comprise a plurality of antibodies, antibody derivatives, or antibody fragments wherein the plurality of such antibody agents binds specifically with a marker protein, or a fragment of the protein. In an additional embodiment, the kit comprises a nucleic acid probe that binds specifically with a marker nucleic acid or a fragment of the nucleic acid. The kit may also comprise a plurality of probes, wherein each of the probes binds specifically with a marker nucleic acid, or a fragment of the nucleic acid.

[0023] In a further aspect, the invention relates to methods for treating a patient afflicted with or at risk of developing RA. Such methods may comprise reducing the expression and/or interfering with the biological function of a marker of the invention. In one embodiment, the method comprises providing to the patient an antisense oligonucleotide or polynucleotide complementary to a marker nucleic acid, or a segment thereof. For example, an antisense polynucleotide may be provided to the patient through the delivery of a vector that expresses an anti-sense polynucleotide of a marker nucleic acid or a fragment thereof. In another embodiment, the method comprises providing to the patient an antibody, an antibody derivative, or antibody fragment, which binds specifically with a marker protein or a fragment of the protein.

[0024] It will be appreciated that the methods and kits of the present invention may also include known RA markers, *i.e.*, the markers of the present invention may be used alone, in combination, and in combination with known RA markers.

DETAILED DESCRIPTION OF THE INVENTION

[0025] The invention relates to newly discovered markers associated with RA. It has been discovered that a higher than normal level of expression of individual markers and combinations of markers described herein correlates with RA. Methods are provided for detecting the presence of RA, the absence of RA, the type of RA (*e.g.*, erosive versus non-erosive), and other characteristics of RA that are relevant to prevention, diagnosis, characterization, and therapy of RA.

Definitions

[0026] As used herein, each of the following terms has the meaning associated with it in this section.

[0027] The articles "a" and "an" are used herein to refer to one or to more than one (*i.e.* to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

[0028] A "marker" is a naturally-occurring polymer corresponding to at least one of the proteins listed in Tables 1 and 2. Markers further include, without limitation, sense and anti-sense strands of genomic DNA (*i.e.* including any introns occurring therein), RNA generated by transcription of genomic DNA (*i.e.* prior to splicing), RNA generated by splicing of RNA transcribed from genomic DNA, and proteins generated by translation of spliced RNA (*e.g.* including proteins both before and after cleavage of normally cleaved regions such as transmembrane signal sequences). As used herein, "marker" may also include a cDNA made by reverse transcription of an RNA generated by transcription of genomic DNA (including spliced RNA).

[0029] A "marker set" is a group of more than one marker.

[0030] "Proteins of the invention" encompass marker proteins and their fragments; variant marker proteins and their fragments; peptides and polypeptides comprising an at least 15 amino acid segment of a marker or variant marker protein; and fusion proteins comprising a marker or variant marker protein, or an at least 15 amino acid segment of a marker or variant marker protein.

[0031] Unless otherwise specified herewithin, the terms "antibody" and "antibodies" broadly encompass naturally-occurring forms of antibodies (*e.g.*, IgG, IgA, IgM, IgE) and recombinant antibodies such as single-chain antibodies, chimeric and humanized antibodies and multi-specific antibodies, as well as fragments and derivatives of all of the foregoing, which fragments and derivatives have at least an antigenic binding site. Antibody derivatives may comprise a protein or chemical moiety conjugated to an antibody.

[0032] As used herein a "polynucleotide corresponds to" another (a first) polynucleotide if it is related to the first polynucleotide by any of the following relationships: 1) The second polynucleotide comprises the first polynucleotide and the second polynucleotide encodes a gene product. 2) The second polynucleotide is 5' or 3' to the first polynucleotide in cDNA, RNA, genomic DNA, or fragment of any of these polynucleotides. For example, a second polynucleotide may be fragment of a gene that includes the first and second polynucleotides. The first and second polynucleotides are related in that they are components of the gene coding for a gene product, such as a protein or antibody. However, it is not necessary that the second polynucleotide comprises or overlaps with the first polynucleotide to be encompassed within the definition of "corresponding to" as used herein. For example, the first polynucleotide may be a fragment of a 3' untranslated region of the second polynucleotide. The first and second polynucleotide may be fragments of a gene coding for a gene product. The second polynucleotide may be an exon of the gene while the first polynucleotide may be an intron of the gene. 3) The second polynucleotide is the complement of the first polynucleotide.

[0033] The term "probe" refers to any molecule which is capable of selectively binding to a specifically intended target molecule, for example a marker of the invention. Probes can be either synthesized by one skilled in the art, or derived from appropriate biological preparations. For purposes of detection of the target molecule, probes may be specifically designed to be labeled, as described herein. Examples of molecules that can be utilized as probes include, but are not limited to, RNA, DNA, proteins, antibodies, and organic monomers.

[0034] An "RA-associated" body fluid or "patient sample" includes, without limitation, blood fluids (*e.g.* whole blood, blood serum, plasma, blood having platelets removed therefrom, etc.), synovial fluid, urine, saliva and tears.

[0035] "Expression" refers to the presence or abundance of a marker protein or a fragment of the protein in a sample as well as the presence of a marker nucleic acid, *i.e.*, a transcribed polynucleotide (*e.g.*, an mRNA or a cDNA), or a fragment thereof, in a sample.

[0036] "Over-expression" and "under-expression" of a marker refers to expression of the marker in a sample, at a greater or lesser level, respectively, than the normal level of expression of the marker (*e.g.* at least two-fold greater or lesser level). The marker is said to be over-expressed or under-expressed if either the marker protein or marker nucleic acid is present at a greater or lesser level, respectively, than the normal level in a patient sample.

[0037] "Erosive RA" is RA characterized by erosions or pits in the surface of the bone adjacent to the articular surface. In particular, in erosive RA, the granulation tissue actively invades and destroys the periarticular bone and cartilage at the margin between the synorium and the bone.

[0038] "Non-erosive RA" is RA that does not exhibit erosive RA characteristics.

[0039] As used herein, the term "promoter/regulatory sequence" means a nucleic acid sequence which is required for expression of a gene product operably linked to the promoter/regulatory sequence. In some instances, this sequence may be the core promoter sequence and in other instances, this sequence may also include an enhancer sequence and other regulatory elements which are required for expression of the gene product. The promoter/regulatory sequence may, for example, be one which expresses the gene product in a tissue-specific manner.

[0040] A "constitutive" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living human cell under most or all physiological conditions of the cell.

[0041] An "inducible" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living human cell substantially only when an inducer, which corresponds to the promoter, is present in the cell.

[0042] A "tissue-specific" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living human cell substantially only if the cell is a cell of the tissue type corresponding to the promoter.

[0043] A "transcribed polynucleotide" is a polynucleotide (*e.g.* an RNA, a cDNA, or an analog of one of an RNA or cDNA) which is complementary to or homologous with all or a portion of a mature RNA made by transcription of a genomic DNA corresponding to a marker of the invention and normal post-transcriptional processing (*e.g.* splicing), if any, of the transcript.

[0044] "Complementary" refers to the broad concept of sequence complementarity between regions of two nucleic acid strands or between two regions of the same nucleic acid strand. It is known that an adenine residue of a first nucleic acid region is capable of forming specific hydrogen bonds ("base pairing") with a residue of a second nucleic acid region which is antiparallel to the first region if the residue is thymine or uracil. Similarly, it is known that a cytosine residue of a first nucleic acid strand is capable of base pairing with a residue of a second nucleic acid strand which is antiparallel to the first strand if the residue is guanine. A first region of a nucleic acid is complementary to a second region of the same or a different nucleic acid if, when the two regions are arranged in an antiparallel fashion, at least one nucleotide residue of the first region is capable of base pairing with a residue of the second region. Preferably, the first region comprises a first portion and the second region comprises a second portion, whereby, when the first and second portions are arranged in an antiparallel fashion, at least about 50%, and preferably at least about 75%, at least about 90%, or at least about 95% of the nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion. More preferably, all nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion.

[0045] "Homologous" as used herein, refers to nucleotide sequence similarity between two regions of the same nucleic acid strand or between regions of two different nucleic acid strands. When a nucleotide residue position in both regions is occupied by the same nucleotide residue, then the regions are homologous at that position. A first region is homologous to a second region if at least one nucleotide residue position of each region is occupied by the same residue. Homology between two regions is expressed in terms of the proportion of nucleotide residue positions of the two regions that are occupied by the same nucleotide residue. By way of example, a region having the nucleotide sequence 5'-ATTGCC-3' and a region having the nucleotide sequence 5'-TATGGC-3' share 50% homology. Preferably, the first region comprises a first portion and the second region comprises a second portion, whereby, at least about 50%, and preferably at least about 75%, at least about 90%, or at least about 95% of the nucleotide residue positions of each

of the portions are occupied by the same nucleotide residue. More preferably, all nucleotide residue positions of each of the portions are occupied by the same nucleotide residue.

[0046] A marker is "fixed" to a substrate if it is covalently or non-covalently associated with the substrate such the substrate can be rinsed with a fluid (*e.g.* standard saline citrate, pH 7.4) without a substantial fraction of the marker dissociating from the substrate.

[0047] As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.* encodes a natural protein).

[0048] The term "isoform" as used herein refers to variants of a polypeptide that are encoded by the same gene, but that differ in their pI or MW, or both. Such isoforms can differ in their amino acid composition (*e.g.*, as a result of alternative mRNA or premRNA processing, *e.g.* alternative splicing or limited proteolysis) and in addition, or in the alternative, may arise from differential post-translational modification (*e.g.*, glycosylation, acylation, phosphorylation).

[0049] Expression of a marker in a patient is "significantly" higher or lower than the normal level of expression of a marker if the level of expression of the marker is greater or less, respectively, than the normal level by an amount greater than the standard error of the assay employed to assess expression, and preferably at least twice, and more preferably three, four, five or ten times that amount. Alternately, expression of the marker in the patient can be considered "significantly" higher or lower than the normal level of expression if the level of expression is at least about two, and preferably at least about three, four, or five times, higher or lower, respectively, than the normal level of expression of the marker.

[0050] RA is "inhibited" if at least one symptom of the RA is alleviated, terminated, slowed, or prevented. As used herein, RA is "inhibited" if recurrence of RA is reduced, slowed, delayed, or prevented or RA remission is induced or maintained.

[0051] A kit is any manufacture (*e.g.* a package or container) comprising at least one reagent, *e.g.* a probe, for specifically detecting a marker of the invention. The manufacture may be promoted, distributed, or sold as a unit for performing the methods of the present invention.

[0052] The present invention is based, in part, on newly identified markers which are differently expressed in RA patients as compared to normal individuals (*i.e.,* individuals not afflicted by RA). The markers of the invention correspond to polypeptide and nucleic acid molecules which can be detected in one or both of normal samples and diseased patient samples. The presence, absence, or level of expression of one or more of these markers in patient samples is herein correlated with the rheumatoid arthritic state of the patient.

[0053] The present invention also provides markers which are differently expressed in patients with erosive RA. Erosive RA is characterized by erosions or pits in the surface of the bone adjacent to the articular surface. In particular, in erosive RA, the granulation tissue actively invades and destroys the periarticular bone and cartilage at the margin between the synovium and the bone.

[0054] The compositions, kits, and methods of the invention have the following uses, among others:

assessing whether a patient is afflicted with RA;

assessing the stage of RA in a patient;

assessing the progressive nature of RA in a patient;

assessing whether a patient has erosive RA;

assessing whether a patient has non-erosive RA;

making an isolated hybridoma which produces an antibody useful for assessing whether a patient is afflicted with RA;

assessing the efficacy of one or more test compounds for inhibiting RA in a patient;

assessing the efficacy of a therapy for inhibiting RA in a patient;

assessing the efficacy of a therapy for inhibiting erosive RA in a patient;

assessing the efficacy of a therapy for inhibiting non-erosive RA in a patient;

monitoring the progression of RA in a patient;

selecting a composition or therapy for inhibiting RA in a patient;

selecting a composition or therapy for inhibiting erosive RA in a patient;

selecting a composition or therapy for inhibiting non-erosive RA in a patient;

developing agents effective in treating synovitis;

developing agents effective in treating erosive RA;

developing agents effective in treating non-erosive RA;

treating a patient afflicted with RA;

inhibiting RA in a patient;

assessing the rhematoid arthritic progressive potential of a test compound; and

inhibiting RA in a patient at risk for developing RA.

[0055] The methods of the present invention comprise the step of comparing the level of expression of a marker in a patient sample, with the normal level of expression of the marker. A significant difference between the level of expression of the marker in the patient sample and the normal level is an indication that the patient is afflicted with RA. A "normal" level of expression refers to the expression level of the marker in the control, such as in a sample from an individual without RA. Subjects that are not afflicted with RA can include normal subjects with no known disease or condition, or subjects with joint diseases or conditions other than RA, including gout, osteoarthritis, or synovitis (*e.g.*, traumatic synovitis). Alternatively, and particularly as further information becomes available as a result of routine performance of the methods described herein, population-average values for expression of the markers of the invention may be used as the "normal" level of expression. For example, a laboratory may establish reference ranges for the level of the marker for subjects with and without RA, as well as for subjects with erosive and non-erosive forms of RA, as is conventional in the diagnostic art.

[0056] As used herein the term "expression" refers to the presence or abundance of a marker protein or a fragment of the protein in a sample as well as the presence of a marker nucleic acid, *i.e.*, a transcribed polynucleotide (*e.g.*, an mRNA or a cDNA), or a fragment thereof, in a sample. In a method of determining the abundance of a marker in a sample compared to a normal or control, *i.e.*, to identify markers that are differentially present, the relative abundance may be determined by normalizing the signal obtained upon detecting the marker in a sample by reference to a suitable background parameter, *e.g.*, to the total protein in the sample being analyzed to an invariant marker, *i.e.*, a marker whose abundance is known to be similar in the sample being compared, or to the total signal detected from all proteins in the sample.

[0057] In a preferred diagnostic method of assessing whether a patient is afflicted with RA (*e.g.*, new detection ("screening") and detection of recurrence), the method comprises comparing:

a) the level of expression of a marker of the invention in a patient sample, and

b) the normal level of expression of the marker in a control.

[0058] A significantly higher level of expression of the marker in the patient sample as compared to the normal level is an indication that the patient is afflicted with RA. In one embodiment, the marker is listed in Table 2.

[0059] In a further preferred diagnostic method of assessing whether a patient is afflicted with erosive RA, the method comprises comparing:

a) the level of expression of a marker of the invention in a patient sample, and

b) the normal level of expression of the marker in a control.

[0060] A significantly higher level of expression of the marker in the patient sample as compared to the normal level is an indication that the patient is afflicted with erosive RA. In one embodiment, the marker is listed in Table 1. In a preferred embodiment, the marker is listed in Table 2.

[0061] The invention also provides diagnostic methods for assessing the efficacy of a therapy for inhibiting RA in a patient. Such methods comprise comparing:

a) expression of a marker of the invention in a first sample obtained from the patient prior to providing at least a portion of the therapy to the patient, and

b) expression of the marker in a second sample obtained from the patient following provision of the portion of the therapy.

[0062] A significantly lower level of expression of the marker in the second sample relative to that in the first sample is an indication that the therapy is efficacious for inhibiting RA in the patient. It will be appreciated that in these methods the "therapy" may be any therapy for treating RA including, but not limited to, anti-inflammatory drugs, disease-modifying drugs and gene therapy. Thus, the methods of the invention may be used to evaluate a patient before, during and after therapy, for example, to evaluate the efficacy of treatment.

[0063] In a preferred embodiment, the methods are directed to therapy using a chemical or biologic agent. These methods comprise comparing:

a) expression of a marker of the invention in a first sample obtained from the patient and maintained in the presence of the chemical or biologic agent, and

b) expression of the marker in a second sample obtained from the patient and maintained in the absence of the agent.

[0064] A significantly lower level of expression of the marker in the first sample relative to that in the second sample is an indication that the agent is efficacious for inhibiting RA in the patient. In one embodiment, the first and second samples can be

portions of a single sample obtained from the patient or portions of pooled samples obtained from the patient.

[0065] The invention additionally provides a monitoring method for assessing the progression of RA in a patient, the method comprising:

a) detecting in a patient sample at a first time point, the expression of a marker of the invention;

b) repeating step a) at a subsequent point in time; and

c) comparing the level of expression detected in steps a) and b), and therefrom monitoring the progression of RA in the patient.

[0066] A significantly higher level of expression of the marker in the sample at the subsequent time point from that of the sample at the first time point is an indication that the RA has progressed, whereas a significantly lower level of expression is an indication that the RA has regressed.

[0067] The invention moreover provides a test method for selecting a composition for inhibiting RA in a patient. This method comprises the steps of:

a) obtaining a sample from the patient;

b) separately maintaining aliquots of the sample in the presence of a plurality of test compositions;

c) comparing expression of a marker of the invention in each of the aliquots; and

d) selecting one of the test compositions which significantly reduces the level of expression of the marker in the aliquot containing that test composition, relative to the levels of expression of the marker in the presence of the other test compositions.

[0068] In addition, the invention further provides a method of inhibiting RA in a patient. This method comprises the steps of:

a) obtaining a sample from the patient;

b) separately maintaining aliquots of the sample in the presence of a plurality of compositions;

c) comparing expression of a marker of the invention in each of the aliquots; and

d) administering to the patient at least one of the compositions which significantly lowers the level of expression of the marker in the aliquot containing that composition, relative to the levels of

expression of the marker in the presence of the other compositions.

[0069] Any marker or combination of markers listed in the tables, as well as any known markers in combination with the markers listed in the tables, may be used in the compositions, kits, and methods of the present invention. In general, it is preferable to use markers for which the difference between the level of expression of the marker in RA patient samples and the level of expression of the same marker in normal samples is as great as possible. Although this difference can be as small as the limit of detection of the method for assessing expression of the marker, it is preferred that the difference be at least greater than the standard error of the assessment method, and preferably a difference of at least 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 15-, 20-, 25-, 100-, 500-, 1000-fold or greater.

[0070] It will be appreciated that patient samples containing bodily fluids (*e.g.*, blood fluid, whole blood, serum, blood having platelets removed therefrom etc., and synovial fluid) may be used in the methods of the present invention. In these embodiments, the level of expression of the marker can be assessed by assessing the amount or abundance (*e.g.* absolute amount or concentration) of a marker product (*e.g.*, protein and RNA transcript encoding said protein, fragments of the protein, isoforms of the protein, and RNA transcript) in a sample. The sample can, of course, be subjected to a variety of well-known post-collection preparative and storage techniques (*e.g.* fixation, storage, freezing, lysis, homogenization, DNA or RNA extraction, ultrafiltration, concentration, evaporation, centrifugation, etc.) prior to assessing the amount of the marker in the sample.

[0071] Preferred *in vivo* techniques for detection of a marker protein of the invention include introducing into a subject an antibody that specifically binds the protein, isoform of the protein, or protein fragment. In certain embodiments, the antibody can be labeled with a radioactive molecule whose presence and location in a subject can be detected by standard imaging techniques.

[0072] Expression of a marker of the invention may be assessed by any of a wide variety of well known methods for detecting expression of a protein or transcribed molecule. Non-limiting examples of such methods include immunological methods for detection of secreted, cell-surface, cytoplasmic, or nuclear proteins, protein purification methods, protein function or activity assays, nucleic acid hybridization methods, nucleic acid reverse transcription methods, and nucleic acid amplification methods. Such methods

may also include physical methods such as liquid and gas chromatography, mass spectroscopy, nuclear magnetic resonance and other imaging technologies.

[0073] In a preferred embodiment, expression of a marker protein is assessed using an antibody (*e.g.* a radio-labeled, chromophore-labeled, fluorophore-labeled, or enzyme-labeled antibody), an antibody derivative (*e.g.* an antibody conjugated with a substrate or with the protein or ligand of a protein-ligand pair {*e.g.* biotin-streptavidin}), or an antibody fragment (*e.g.* a single-chain antibody, an isolated antibody hypervariable domain, etc.) which binds specifically with a marker protein, isoform of the marker protein, or a fragment of the protein, wherein the protein may have undergone none, all or a portion of its normal post-translational modification and/or proteolysis during the course of its secretion or release from cells.

[0074] In another preferred embodiment, expression of a marker is assessed by preparing mRNA/cDNA (*i.e.* a transcribed polynucleotide) from cells in a patient sample, and by hybridizing the mRNA/cDNA with a reference polynucleotide which comprises the marker nucleic acid sequence or its complement, or a fragment of said sequence or complement. cDNA can, optionally, be amplified using any of a variety of polymerase chain reaction methods prior to hybridization with the reference polynucleotide. Expression of one or more marker nucleic acid can likewise be detected using quantitative PCR to assess the level of RNA transcripts encoded by the marker(s).

[0075] In a related embodiment, a mixture of transcribed polynucleotides obtained from the sample is contacted with a substrate having fixed thereto a polynucleotide complementary to or homologous with at least a portion (*e.g.* at least 7, 10, 15, 20, 25, 30, 40, 50, 100, 500, or more nucleotide residues) of a RNA transcript encoded by a marker of the invention. If polynucleotides complementary to or homologous with a RNA transcript encoded by the marker of the invention are differentially detectable on the substrate (*e.g.* detectable using radioactivity, different chromophores or fluorophores), are fixed to different selected positions, then the levels of expression of a plurality of markers can be assessed simultaneously using a single substrate (*e.g.* a "gene chip" microarray of polynucleotides fixed at selected positions). When a method of assessing marker expression is used which involves hybridization of one nucleic acid with another, it is preferred that the hybridization be performed under stringent hybridization conditions.

[0076] Because the compositions, kits, and methods of the invention rely on detection of a difference in expression levels of one or more markers of the invention, it is preferable that the level of expression of the marker is significantly greater than the

minimum detection limit of the method used to assess expression in a normal or control sample.

[0077] It is understood that by routine screening of additional patient samples for the expression levels of one or more of the markers of the invention, it will be realized that certain of the markers are expressed at varying levels based on the progressiveness of disease. Thus the markers and methods of the present invention may be used to identify a non-progressive to progressive gradient. Such gradient would be especially useful in characterizing, managing and treating RA.

[0078] It is recognized that certain markers correspond to proteins which are secreted from patient samples (*i.e.* synovial fluid, endothelial cells, synovium cells, serum, plasma) to the extracellular space surrounding the cells. These markers are preferably used in certain embodiments of the compositions, kits, and methods of the invention, owing to the fact that the protein corresponding to each of these markers can be detected in an RA-associated body fluid sample, which may be easily collected from a human patient. It will be appreciated, however, that intracellular markers are also included within the markers of the present invention and are also useful in the methods of the present invention.

[0079] It is a simple matter for the skilled artisan to determine whether any particular marker corresponds to a secreted protein. In order to make this determination, the protein corresponding to a marker is expressed in a test cell, extracellular fluid is collected, and the presence or absence of the protein in the extracellular fluid is assessed (*e.g.* using a labeled antibody which binds specifically with the protein).

[0080] The compositions, kits, and methods of the invention can also be used to detect expression of markers corresponding to proteins having at least one portion which is displayed on the surface of cells which express it. It is a simple matter for the skilled artisan to determine whether the protein corresponding to any particular marker comprises a cell-surface protein. For example, immunological methods may be used to detect such proteins on whole cells, or well known computer-based sequence analysis methods (*e.g.* the SIGNALP program; Nielsen *et al.*, 1997, *Protein Engineering* 10:1-6) may be used to predict the presence of at least one extracellular domain (*i.e.* including both secreted proteins and proteins having at least one cell-surface domain). Expression of a marker corresponding to a protein having at least one portion which is displayed on the surface of a cell which expresses it may be detected without necessarily lysing the cell (*e.g.* using a labeled antibody which binds specifically with a cell-surface domain of the protein).

[0081] When a plurality of markers of the invention are used in the compositions, kits, and methods of the invention, the level of expression of each marker in a patient sample can be compared with the normal level of expression of each of the plurality of markers in RA samples of the same type, either in a single reaction mixture (*i.e.* using reagents, such as different fluorescent probes, for each marker) or in individual reaction mixtures corresponding to one or more of the markers. In one embodiment, a significantly enhanced level of expression of more than one of the plurality of markers in the sample, relative to the corresponding normal levels, is an indication that the patient is afflicted with RA. When a plurality of markers is used, it is preferred that 2, 3, 4, 5, 8, 10, 12, 15, 20, 30, or 50 or more individual markers be used, wherein fewer markers are preferred.

[0082] Prior to the present invention, only a limited number of markers were known to be associated with RA (*e.g.*, RF, complement factor B, and C-reactive protein). These markers may be used together with one or more markers of the invention in a panel of markers. For example, a sample may be assayed to determine the presence and/or expression levels of known markers in combination with the markers of the present invention. The presence, over- and/or under-expression of markers, such as RF in combination with the presence, over- and/or underexpression of the markers of the present invention, may be used to further characterize RA.

[0083] It is recognized that the compositions, kits, and methods of the invention will be of particular utility to patients having an enhanced risk of developing RA and their medical advisors. Patients recognized as having an enhanced risk of developing RA include, for example, patients having a familial history of RA, patients identified as having a RF, patients of advancing age and women of advancing age (*i.e.* between 40 and 60 years).

[0084] The level of expression of a marker in normal (*i.e.* an individual who is not afflicted with RA) individuals or a control can be assessed in a variety of ways. As further information becomes available as a result of routine performance of the methods described herein, population-average values for expression of the markers of the invention may be used. In other embodiments, the 'normal' level of expression of a marker may be determined by assessing expression of the marker in a patient sample obtained from a non-RA-afflicted patient, from a patient sample obtained from a patient before the suspected onset of RA in the patient, from archived patient samples, and the like.

[0085] The invention includes compositions, kits, and methods for assessing the presence of RA in a sample (*e.g.* an archived tissue sample or a sample obtained from a

patient). These compositions, kits, and methods are substantially the same as those described above, except that, where necessary, the compositions, kits, and methods are adapted for use with samples other than patient samples. For example, when the sample to be used is a parafinized, archived human tissue sample, it can be necessary to adjust the ratio of compounds in the compositions of the invention, in the kits of the invention, or the methods used to assess levels of marker expression in the sample. Such methods are well known in the art and within the skill of the ordinary artisan.

[0086] The invention includes a kit for assessing the presence of RA (*e.g.* in a sample such as a patient sample). The kit comprises a plurality of reagents, each of which is capable of binding specifically with a nucleic acid or polypeptide corresponding to a marker of the invention. Suitable reagents for binding with a polypeptide corresponding to a marker of the invention include antibodies, antibody derivatives, antibody fragments, and the like. Suitable reagents for binding with a nucleic acid (*e.g.* a genomic DNA, an mRNA, a spliced mRNA, a cDNA, or the like) include complementary nucleic acids. For example, the nucleic acid reagents may include oligonucleotides (labeled or non-labeled) fixed to a substrate, labeled oligonucleotides not bound with a substrate, pairs of PCR primers, molecular beacon probes, and the like.

[0087] The kit of the invention may optionally comprise additional components useful for performing the methods of the invention. By way of example, the kit may comprise fluids (*e.g.* SSC buffer) suitable for annealing complementary nucleic acids or for binding an antibody with a protein with which it specifically binds, one or more sample compartments, an instructional material which describes performance of a method of the invention, a sample from a normal individual, a sample from a RA patient, and the like.

[0088] The invention also includes a method of making an isolated hybridoma which produces an antibody useful for assessing whether patient is afflicted with RA. In this method, a marker protein of the invention is isolated (*e.g.* by purification from a cell in which it is expressed or by transcription and translation of a nucleic acid encoding the protein *in vivo* or *in vitro* using known methods). A vertebrate, preferably a mammal such as a mouse, rat, rabbit, or sheep, is immunized using the isolated protein or protein fragment. The vertebrate may optionally (and preferably) be immunized at least one additional time with the isolated protein or protein fragment, so that the vertebrate exhibits a robust immune response to the protein or protein fragment. Splenocytes are isolated from the immunized vertebrate and fused with an immortalized cell line to form hybridomas, using any of a variety of methods well known in the art. Hybridomas formed

in this manner are then screened using standard methods to identify one or more hybridomas which produce an antibody which specifically binds with the protein or protein fragment. The invention also includes hybridomas made by this method and antibodies made using such hybridomas.

[0089] The invention also includes a method of assessing the efficacy of a test compound for inhibiting RA. As described above, differences in the level of expression of the markers of the invention correlate with the rheumatoid arthritic state of the patient. Although it is recognized that changes in the levels of expression of certain of the markers of the invention likely result from the rheumatoid arthritic state of patient, it is likewise recognized that changes in the levels of expression of other of the markers of the invention induce, maintain, and promote the rheumatoid arthritic state of those patients. Thus, compounds which inhibit RA in a patient will cause the level of expression of one or more of the markers of the invention to change to a level nearer the normal level of expression for that marker (*i.e.* the level of expression for the marker in RA patients).

[0090] This method thus comprises comparing expression of a marker in a first patient sample and maintained in the presence of the test compound and expression of the marker in a second patient sample and maintained in the absence of the test compound. A significant decrease in the level of expression of a marker may be an indication that the test compound inhibits RA. The patient samples may, for example, be aliquots of a single sample obtained from a patient, pooled normal samples obtained an individual, cells of a normal individual, aliquots of a single sample obtained from a RA patient, pooled samples from a RA patient, cells of a RA patient, or the like. In one embodiment, the samples from a RA patient and a plurality of compounds known to be effective for inhibiting RA are tested in order to identify the compound which is likely to best inhibit the RA in the patient.

[0091] This method may likewise be used to assess the efficacy of a therapy for inhibiting RA in a patient. In this method, the level of expression of one or more markers of the invention in a pair of samples (one subjected to the therapy, the other not subjected to the therapy) is assessed. As with the method of assessing the efficacy of test compounds, if the therapy induces a significant decrease in the level of expression of a marker, or blocks induction of a marker, then the therapy may be efficacious for inhibiting RA. As above, if samples from a selected patient are used in this method, then alternative therapies can be assessed *in vitro* in order to select a therapy most likely to be efficacious for inhibiting RA in the patient.

[0092] The present invention further provides methods for identifying the presence of erosive and non-erosive RA by detecting expression of a marker listed in Tables 1 and 2, wherein over-expression of one or a plurality of the markers is correlated with erosive RA. By identifying whether a patient sample is afflicted with erosive or non-erosive RA, therapy may be customized to better treat the specific type of RA.

[0093] Expression of a marker can be inhibited in a number of ways generally known in the art. For example, an antisense oligonucleotide can be provided to the patient samples in order to inhibit transcription, translation, or both, of the marker(s). Alternately, a polynucleotide encoding an antibody, an antibody derivative, or an antibody fragment, and operably linked with an appropriate promoter/regulator region, can be provided to the patient sample in order to generate intracellular antibodies which will inhibit the function or activity of the protein. Using the methods described herein, a variety of molecules, particularly including molecules sufficiently small that they are able to cross the cell membrane, can be screened in order to identify molecules which inhibit expression of the marker(s). The compound so identified can be provided to the patient in order to inhibit expression of the marker(s) in the patient.

[0094] Expression of a marker can be enhanced in a number of ways generally known in the art. For example, a polynucleotide encoding the marker and operably linked with an appropriate promoter/regulator region can be provided to patient samples in order to induce enhanced expression of the protein (and mRNA) corresponding to the marker therein. Alternatively, if the protein is capable of crossing the cell membrane, inserting itself in the cell membrane, or is normally a secreted protein, then expression of the protein can be enhanced by providing the protein (*e.g.* directly or by way of the bloodstream) to the patient sample.

[0095] As described above, the rheumatoid arthritic state of the patient is correlated with changes in the levels of expression of the markers of the invention. The invention thus includes a method for assessing the RA promoting or progression potential of a test compound. This method comprises maintaining separate aliquots of patient samples in the presence and absence of the test compound. Expression of a marker of the invention in each of the aliquots is compared. A significant increase in the level of expression of a marker in the aliquot maintained in the presence of the test compound (relative to the aliquot maintained in the absence of the test compound) may be an indication that the test compound possesses RA promoting or progression potential. The relative RA promoting or progression potentials of various test compounds can be assessed

by comparing the degree of enhancement or inhibition of the level of expression of the relevant markers, by comparing the number of markers for which the level of expression is enhanced or inhibited, or by comparing both.

[0096] Various aspects of the invention are described in further detail in the following subsections.

I. Isolated Proteins and Antibodies

[0097] One aspect of the invention pertains to marker proteins which are isolated proteins biologically active portions thereof, isoforms, as well as polypeptide fragments suitable for use as immunogens to raise antibodies directed against a polypeptide of the invention. In one embodiment, the native polypeptide corresponding to a marker can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, polypeptides corresponding to a marker of the invention are produced by recombinant DNA techniques. Alternative to recombinant expression, a polypeptide corresponding to a marker of the invention can be synthesized chemically using standard peptide synthesis techniques.

[0098] An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free of chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, protein that is substantially free of cellular material includes preparations of protein having less than about 30%, 20%, 10%, or 5% (by dry weight) of heterologous protein (also referred to herein as a "contaminating protein"). When the protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, 10%, or 5% of the volume of the protein preparation. When the protein is produced by chemical synthesis, it is preferably substantially free of chemical precursors or other chemicals, *i.e.*, it is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. Accordingly such preparations of the protein have less than about 30%, 20%, 10%, 5% (by dry weight) of chemical precursors or compounds other than the polypeptide of interest.

[0099] Biologically active portions of a polypeptide corresponding to a marker of the invention include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the protein corresponding to the marker, which include fewer amino acids than the full length protein, and exhibit at least one activity of the corresponding full-length protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the corresponding protein. A biologically active portion of a protein of the invention can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of the native form of a polypeptide of the invention.

[00100] Preferred polypeptides have amino acid sequences encoded by the nucleic acid sequences described herein. Other useful proteins are substantially identical (*e.g.*, at least about 40%, preferably 50%, 60%, 70%, 80%, 90%, 95%, or 99%) to one of these sequences and retain the functional activity of the protein of the corresponding naturally-occurring protein yet differ in amino acid sequence due to natural allelic variation or mutagenesis.

[00101] To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = # of identical positions/total # of positions (*e.g.*, overlapping positions) x100). In one embodiment the two sequences are the same length.

[00102] The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, *et al.* (1990) *J. Mol.*

Biol. 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Yet another useful algorithm for identifying regions of local sequence similarity and alignment is the FASTA algorithm as described in Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448. When using the FASTA algorithm for comparing nucleotide or amino acid sequences, a PAM120 weight residue table can, for example, be used with a *k*-tuple value of 2.

[00103] The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, only exact matches are counted.

[00104] The invention also provides chimeric or fusion proteins corresponding to a marker of the invention. As used herein, a "chimeric protein" or "fusion protein" comprises all or part (preferably a biologically active part) of a polypeptide corresponding to a marker of the invention operably linked to a heterologous polypeptide (*i.e.*, a polypeptide other than the polypeptide corresponding to the marker). Within the fusion protein, the term "operably linked" is intended to indicate that the polypeptide of the invention and the heterologous polypeptide are fused in-frame to each other. The heterologous polypeptide can be fused to the amino-terminus or the carboxyl-terminus of the polypeptide of the invention.

[00105] One useful fusion protein is a GST fusion protein in which a polypeptide corresponding to a marker of the invention is fused to the carboxyl terminus of GST

sequences. Such fusion proteins can facilitate the purification of a recombinant polypeptide of the invention.

[00106] In another embodiment, the fusion protein contains a heterologous signal sequence at its amino terminus. For example, the native signal sequence of a polypeptide corresponding to a marker of the invention can be removed and replaced with a signal sequence from another protein. For example, the gp67 secretory sequence of the baculovirus envelope protein can be used as a heterologous signal sequence (Ausubel *et al.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, NY, 1992). Other examples of eukaryotic heterologous signal sequences include the secretory sequences of melittin and human placental alkaline phosphatase (Stratagene; La Jolla, California). In yet another example, useful prokaryotic heterologous signal sequences include the phoA secretory signal (Sambrook *et al., supra*) and the protein A secretory signal (Pharmacia Biotech; Piscataway, New Jersey).

[00107] In yet another embodiment, the fusion protein is an immunoglobulin fusion protein in which all or part of a polypeptide corresponding to a marker of the invention is fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand (soluble or membrane-bound) and a protein on the surface of a cell (receptor), to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion protein can be used to affect the bioavailability of a cognate ligand of a polypeptide of the invention. Inhibition of ligand/receptor interaction can be useful therapeutically, both for treating proliferative and differentiative disorders and for modulating (*e.g.* promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies directed against a polypeptide of the invention in a subject, to purify ligands and in screening assays to identify molecules which inhibit the interaction of receptors with ligands.

[00108] Chimeric and fusion proteins of the invention can be produced by standard recombinant DNA techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and re-amplified to generate a chimeric gene sequence (see, *e.g.*, Ausubel *et al., supra*). Moreover, many expression vectors are commercially available

that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the polypeptide of the invention.

[00109] A signal sequence can be used to facilitate secretion and isolation of the secreted protein or other proteins of interest. Signal sequences are typically characterized by a core of hydrophobic amino acids which are generally cleaved from the mature protein during secretion in one or more cleavage events. Such signal peptides contain processing sites that allow cleavage of the signal sequence from the mature proteins as they pass through the secretory pathway. Thus, the invention pertains to the described polypeptides having a signal sequence, as well as to polypeptides from which the signal sequence has been proteolytically cleaved (*i.e.*, the cleavage products). In one embodiment, a nucleic acid sequence encoding a signal sequence can be operably linked in an expression vector to a protein of interest, such as a protein which is ordinarily not secreted or is otherwise difficult to isolate. The signal sequence directs secretion of the protein, such as from a eukaryotic host into which the expression vector is transformed, and the signal sequence is subsequently or concurrently cleaved. The protein can then be readily purified from the extracellular medium by art recognized methods. Alternatively, the signal sequence can be linked to the protein of interest using a sequence which facilitates purification, such as with a GST domain.

[00110] It will be appreciated that as an alternative to recombinant expression, the marker proteins of the present invention may be chemically synthesized using standard peptide synthesis techniques.

[00111] The present invention also pertains to variants of the polypeptides corresponding to individual markers of the invention. Such variants have an altered amino acid sequence, *e.g.*, amino acid substitutions or insertions can be made using naturally occurring or non-naturally occurring amino acids, including L- and D-amino acids. Such variants can function as either agonists (mimetics) or as antagonists. Variants can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation. An agonist can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of the protein. An antagonist of a protein can inhibit one or more of the activities of the naturally occurring form of the protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the protein of interest. Thus, specific biological effects can be elicited by treatment with a variant of limited function. Treatment of a subject with a variant having a

subset of the biological activities of the naturally occurring form of the protein can have fewer side effects in a subject relative to treatment with the naturally occurring form of the protein.

[00112] Variants of a protein of the invention which function as either agonists (mimetics) or as antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of the protein of the invention for agonist or antagonist activity. In one embodiment, a variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential protein sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display). There are a variety of methods which can be used to produce libraries of potential variants of the polypeptides of the invention from a degenerate oligonucleotide sequence. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, 1983, *Tetrahedron* 39:3; Itakura *et al.*, 1984, *Annu. Rev. Biochem.* 53:323; Itakura *et al.*, 1984, *Science* 198:1056; Ike *et al.*, 1983 *Nucleic Acid Res.* 11:477).

[00113] In addition, libraries of fragments of the coding sequence of a polypeptide corresponding to a marker of the invention can be used to generate a variegated population of polypeptides for screening and subsequent selection of variants. For example, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of the coding sequence of interest with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes amino terminal and internal fragments of various sizes of the protein of interest.

[00114] Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial

genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify variants of a protein of the invention (Arkin and Yourvan, 1992, *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.*, 1993, *Protein Engineering* 6(3):327- 331).

[00115] The present invention also pertains to human orthologs for any non-human nucleic acid or amino acid sequences. The identification of such human orthologs may be determined through conventional Molecular Biology techniques known to someone of ordinary skill in the art, such as blast analysis or library screening, as discussed throughout.

[00116] An isolated polypeptide corresponding to a marker of the invention, or a fragment thereof, can be used as an immunogen to generate antibodies using standard techniques for polyclonal and monoclonal antibody preparation. The full-length polypeptide or protein can be used or, alternatively, the invention provides antigenic peptide fragments for use as immunogens. The antigenic peptide of a protein of the invention comprises at least 8 (preferably 10, 15, 20, or 30 or more) amino acid residues of the amino acid sequence of one of the polypeptides of the invention, and encompasses an epitope of the protein such that an antibody raised against the peptide forms a specific immune complex with a marker of the invention to which the protein corresponds. Preferred epitopes encompassed by the antigenic peptide are regions that are located on the surface of the protein, *e.g.*, hydrophilic regions. Hydrophobicity sequence analysis, hydrophilicity sequence analysis, or similar analyses can be used to identify hydrophilic regions.

[00117] An immunogen typically is used to prepare antibodies by immunizing a suitable (*i.e.* immunocompetent) subject such as a rabbit, goat, mouse, or other mammal or vertebrate. An appropriate immunogenic preparation can contain, for example, recombinantly-expressed or chemically-synthesized polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or a similar immunostimulatory agent.

[00118] Accordingly, another aspect of the invention pertains to antibodies directed against a polypeptide of the invention. The terms "antibody" and "antibody substance" as used interchangeably herein refer to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen

binding site which specifically binds an antigen, such as a polypeptide of the invention, *e.g.*, an epitope of a polypeptide of the invention. A molecule which specifically binds to a given polypeptide of the invention is a molecule which binds the polypeptide, but does not substantially bind other molecules in a sample, *e.g.*, a biological sample, which naturally contains the polypeptide. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope.

[00119] Polyclonal antibodies can be prepared as described above by immunizing a suitable subject with a polypeptide of the invention as an immunogen. Preferred polyclonal antibody compositions are ones that have been selected for antibodies directed against a polypeptide or polypeptides of the invention. Particularly preferred polyclonal antibody preparations are ones that contain only antibodies directed against a polypeptide or polypeptides of the invention. Particularly preferred immunogen compositions are those that contain no other human proteins such as, for example, immunogen compositions made using a non-human host cell for recombinant expression of a polypeptide of the invention. In such a manner, the only human epitope or epitopes recognized by the resulting antibody compositions raised against this immunogen will be present as part of a polypeptide or polypeptides of the invention.

[00120] The antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized polypeptide. If desired, the antibody molecules can be harvested or isolated from the subject (*e.g.*, from the blood, plasma, or serum of the subject) and further purified by well-known techniques, such as protein A chromatography to obtain the IgG fraction. Alternatively, antibodies specific for a protein or polypeptide of the invention can be selected or (*e.g.*, partially purified) or purified by, *e.g.*, affinity chromatography. For example, a recombinantly expressed and purified (or partially purified) protein of the invention is produced as described herein, and covalently or non-covalently coupled to a solid support such as, for example, a chromatography column. The column can then be used to affinity purify antibodies specific for the proteins of the invention from a sample containing antibodies directed against a large number of different epitopes, thereby

generating a substantially purified antibody composition, *i.e.*, one that is substantially free of contaminating antibodies. By a substantially purified antibody composition is meant, in this context, that the antibody sample contains at most only 30% (by dry weight) of contaminating antibodies directed against epitopes other than those of the desired protein or polypeptide of the invention, and preferably at most 20%, yet more preferably at most 10%, and most preferably at most 5% (by dry weight) of the sample is contaminating antibodies. A purified antibody composition means that at least 99% of the antibodies in the composition are directed against the desired protein or polypeptide of the invention.

[00121] At an appropriate time after immunization, *e.g.*, when the specific antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497, the human B cell hybridoma technique (see Kozbor *et al.*, 1983, *Immunol. Today* 4:72), the EBV-hybridoma technique (see Cole *et al.*, pp. 77-96 In *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., 1985) or trioma techniques. The technology for producing hybridomas is well known (see generally *Current Protocols in Immunology*, Coligan *et al.* ed., John Wiley & Sons, New York, 1994). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind the polypeptide of interest, *e.g.*, using a standard ELISA assay.

[00122] Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal antibody directed against a polypeptide of the invention can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with the polypeptide of interest. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223,409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication No. WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275- 1281; Griffiths *et al.* (1993) *EMBO J.* 12:725-734.

[00123] Additionally, recombinant antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. (See, *e.g.,* Cabilly *et al.*, U.S. Patent No. 4,816,567; and Boss *et al.*, U.S. Patent No. 4,816,397, which are incorporated herein by reference in their entirety.) Humanized antibodies are antibody molecules from non-human species having one or more complementarily determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule. (See, *e.g.,* Queen, U.S. Patent No. 5,585,089, which is incorporated herein by reference in its entirety.) Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT Publication No. WO 87/02671; European Patent Application 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No. 4,816,567; European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521- 3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Cancer Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559); Morrison (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *Bio/Techniques* 4:214; U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

[00124] Antibodies of the invention may be used as therapeutic agents in treating RA. In a preferred embodiment, completely human antibodies of the invention are used for therapeutic treatment of human RA patients, particularly those having erosive and non-erosive RA. Such antibodies can be produced, for example, using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. The transgenic mice are immunized in the normal fashion with a selected antigen, *e.g.*, all or a portion of a polypeptide corresponding to a marker of the invention. Monoclonal antibodies directed against the antigen can be obtained using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus,

using such a technique, it is possible to produce therapeutically useful IgG, IgA and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar (1995) *Int. Rev. Immunol.* 13:65-93). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, *e.g.*, U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

[00125] Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, *e.g.*, a murine antibody, is used to guide the selection of a completely human antibody recognizing the same epitope (Jespers *et al.*, 1994, *Bio/technology* 12:899-903).

[00126] An antibody directed against a polypeptide corresponding to a marker of the invention (*e.g.*, a monoclonal antibody) can be used to isolate the polypeptide by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, such an antibody can be used to detect the marker (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the level and pattern of expression of the marker. The antibodies can also be used diagnostically to monitor protein levels in tissues or body fluids (*e.g.* in an ovary-associated body fluid) as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

[00127] Further, an antibody (or fragment thereof) can be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

[00128] The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, .alpha.-interferon, .beta.-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophase colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

[00129] Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon *et al.*, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld *et al.* (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom *et al.*, "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson *et al.* (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera *et al.* (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The

Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin *et al.* (eds.), pp. 303-16 (Academic Press 1985), and Thorpe *et al.*, "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982).

[00130] Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

[00131] Accordingly, in one aspect, the invention provides substantially purified antibodies or fragments thereof, and non-human antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequences of the present invention, an amino acid sequence encoded by the cDNA of the present invention, a fragment of at least 15 amino acid residues of an amino acid sequence of the present invention, an amino acid sequence which is at least 95% identical to the amino acid sequence of the present invention (wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4) and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule consisting of the nucleic acid molecules of the present invention, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. In various embodiments, the substantially purified antibodies of the invention, or fragments thereof, can be human, non-human, chimeric and/or humanized antibodies.

[00132] In another aspect, the invention provides non-human antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of: the amino acid sequence of the present invention, an amino acid sequence encoded by the cDNA of the present invention, a fragment of at least 15 amino acid residues of the amino acid sequence of the present invention, an amino acid sequence which is at least 95% identical to the amino acid sequence of the present invention (wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4) and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule consisting of the nucleic acid molecules of the present invention, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. Such non-human antibodies can be goat, mouse, sheep, horse, chicken,

rabbit, or rat antibodies. Alternatively, the non-human antibodies of the invention can be chimeric and/or humanized antibodies. In addition, the non-human antibodies of the invention can be polyclonal antibodies or monoclonal antibodies.

[00133] In still a further aspect, the invention provides monoclonal antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequences of the present invention, an amino acid sequence encoded by the cDNA of the present invention, a fragment of at least 15 amino acid residues of an amino acid sequence of the present invention, an amino acid sequence which is at least 95% identical to an amino acid sequence of the present invention (wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4) and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule consisting of the nucleic acid molecules of the present invention, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. The monoclonal antibodies can be human, humanized, chimeric and/or non-human antibodies.

[00134] The substantially purified antibodies or fragments thereof may specifically bind to a signal peptide, a secreted sequence, an extracellular domain, a transmembrane or a cytoplasmic domain or cytoplasmic membrane of a polypeptide of the invention. In a particularly preferred embodiment, the substantially purified antibodies or fragments thereof, the non-human antibodies or fragments thereof, and/or the monoclonal antibodies or fragments thereof, of the invention specifically bind to a secreted sequence or an extracellular domain of the amino acid sequences of the present invention.

[00135] Any of the antibodies of the invention can be conjugated to a therapeutic moiety or to a detectable substance. Non-limiting examples of detectable substances that can be conjugated to the antibodies of the invention are an enzyme, a prosthetic group, a fluorescent material, a luminescent material, a bioluminescent material, and a radioactive material.

[00136] The invention also provides a kit containing an antibody of the invention conjugated to a detectable substance, and instructions for use. Still another aspect of the invention is a pharmaceutical composition comprising an antibody of the invention and a pharmaceutically acceptable carrier. In preferred embodiments, the pharmaceutical

composition contains an antibody of the invention, a therapeutic moiety, and a pharmaceutically acceptable carrier.

[00137] Still another aspect of the invention is a method of making an antibody that specifically recognizes a polypeptide of the present invention, the method comprising immunizing a mammal with a polypeptide. The polypeptide used as an immunogen comprises an amino acid sequence selected from the group consisting of the amino acid sequence of the present invention, an amino acid sequence encoded by the cDNA of the nucleic acid molecules of the present invention, a fragment of at least 15 amino acid residues of the amino acid sequence of the present invention, an amino acid sequence which is at least 95% identical to the amino acid sequence of the present invention (wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4) and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule consisting of the nucleic acid molecules of the present invention, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C.

[00138] After immunization, a sample is collected from the mammal that contains an antibody that specifically recognizes the polypeptide. Preferably, the polypeptide is recombinantly produced using a non-human host cell. Optionally, the antibodies can be further purified from the sample using techniques well known to those of skill in the art. The method can further comprise producing a monoclonal antibody- producing cell from the cells of the mammal. Optionally, antibodies are collected from the antibody-producing cell.

[00139] Isolated Nucleic Acid Molecules

[00140] Another aspect of the invention pertains to isolated nucleic acid molecules that correspond to a marker of the invention, including nucleic acids which encode a marker protein of the invention or a portion of such a polypeptide. Isolated nucleic acids of the invention also include nucleic acid molecules sufficient for use as hybridization probes to identify nucleic acid molecules that correspond to a marker of the invention, including nucleic acids which encode a polypeptide corresponding to a marker of the invention, and fragments of such nucleic acid molecules, *e.g.*, those suitable for use as PCR primers for the amplification or mutation of nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or

genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

[00141] An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid molecule. Preferably, an "isolated" nucleic acid molecule is free of sequences (preferably protein-encoding sequences) which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kB, 4 kB, 3 kB, 2 kB, 1 kB, 0.5 kB or 0.1 kB of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

[00142] A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid encoding a marker protein can be isolated using standard molecular biology techniques and the sequence information in the database records described herein. Using all or a portion of such nucleic acid sequences, nucleic acid molecules of the invention can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, ed., *Molecular Cloning: A Laboratory Manual, 2nd ed.,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

[00143] A process for identifying the full-length coding sequence of a marker of the present invention is thus also provided. Any conventional recombinant DNA techniques applicable for isolating polynucleotides may also be employed. One such method involves the 5'-RACE-PCR technique, in which the poly-A mRNA that contains the coding sequence of particular interest is first reverse transcribed with a 3'-primer comprising a sequence disclosed herein. The newly synthesized cDNA strand is then tagged with an anchor primer with a known sequence, which preferably contains a convenient cloning restriction site attached at the 5'end. The tagged cDNA is then amplified with the 3'-primer (or a nested primer sharing sequence homology to the internal sequences of the coding region) and the 5'-anchor primer. The amplification may be conducted under conditions of various levels of stringency to optimize the amplification specificity. 5'-

RACE-PCR can be readily performed using commercial kits (available from, *e.g.*, BRL Life Technologies Inc., Clontech) according to the manufacturer's instructions.

[00144] Isolating the complete coding sequence of a gene can also be carried out in a hybridization assay using a suitable probe. The probe preferably comprises at least 10 nucleotides, and more preferably exhibits sequence homology to the polynucleotides of the markers of the present invention. Other high throughput screens for cDNAs, such as those involving gene chip technology, can also be employed in obtaining the complete cDNA sequence.

[00145] In addition, databases exist that reduce the complexity of ESTs by assembling contiguous EST sequences into tentative genes. For example, TIGR has assembled human ESTs into a datable called THC for tentative human consensus sequences. The THC database allows for a more definitive assignment compared to ESTs alone. Software programs exist (TIGR assembler and TIGEM EST assembly machine and contig assembly program (see Huang, X., 1996, *Genomes* 33:21-23)) that allow for assembling ESTs into contiguous sequences from any organism.

[00146] Alternatively, mRNA from a sample preparation is used to construct cDNA library in the ZAP Express vector following the procedure described in Velculescu *et al.*, 1997, *Science* 270:484. The ZAP Express cDNA synthesis kit (Stratagene) is used accordingly to the manufacturer's protocol. Plates containing 250 to 2000 plaques are hybridized as described in Rupert *et al.*, 1988, *Mol. Cell. Bio.* 8:3104 to oligonucleotide probes with the same conditions previously described for standard probes except that the hybridization temperature is reduced to a room temperature. Washes are performed in 6X standard-saline-citrate 0.1% SDS for 30 minutes at room temperature. The probes are labeled with ^{32}P-ATP trough use of T4 polynucleotide kinase.

[00147] A partial cDNA (3' fragment) can be isolated by 3' directed PCR reaction. This procedure is a modification of the protocol described in Polyak *et al.*, 1997, *Nature* 389:300. Briefly, the procedure uses SAGE tags in PCR reaction such that the resultant PCR product contains the SAGE tag of interest as well as additional cDNA, the length of which is defined by the position of the tag with respect to the 3' end of the cDNA. The cDNA product derived from such a transcript driven PCR reaction can be used for many applications.

[00148] RNA from a source to express the cDNA corresponding to a given tag is first converted to double-stranded cDNA using any standard cDNA protocol. Similar conditions used to generate cDNA for SAGE library construction can be employed except

that a modified oligo-dT primer is used to derive the first strand synthesis. For example, the oligonucleotide of composition 5'-**B**-TCC GGC GCG CCG TTT TCC CAG TCA CGA(30)- 3', contains a poly-T stretch at the 3' end for hybridization and priming from poly-A tails, an M13 priming site for use in subsequent PCR steps, a 5' Biotin label (B) for capture to strepavidin-coated magnetic beads, and an AscI restriction endonuclease site for releasing the cDNA from the strepavidin-coated magnetic beads. Theoretically, any sufficiently-sized DNA region capable of hybridizing to a PCR primer can be used as well as any other 8 base pair recognizing endonuclease.

[00149] cDNA constructed utilizing this or similar modified oligo-dT primer is then processed exactly as described in U.S. Patent No. 5,695,937 up until adapter ligation where only one adapter is ligated to the cDNA pool. After Adapter ligation, the cDNA is released from the streptavidin-coated magnetic beads and is then used as a template for cDNA amplification.

[00150] Various PCR protocols can be employed using PCR priming sites within the 3' modified oligo-dT primer and the SAGE tag. The SAGE tag-derived PCR primer employed can be of varying length dictated by 5' extension of the tag into the adaptor sequence. cDNA products are now available for a variety of applications.

[00151] This technique can be further modified by: (1) altering the length and/or content of the modified oligo-dT primer; (2) ligating adaptors other than that previously employed within the SAGE protocol; (3) performing PCR from template retained on the streptavidin-coated magnetic beads; and (4) priming first strand cDNA synthesis with non-oligo-dT based primers.

[00152] Gene trapper technology can also be used. The reagents and manufacturer's instructions for this technology are commercially available from Life Technologies, Inc., Gaithsburg, Maryland. Briefly, a complex population of single-stranded phagemid DNA containing directional cDNA inserts is enriched for the target sequence by hybridization in solution to a biotinylated oligonucleotide probe complementary to the target sequence. The hybrids are captured on streptavidin-coated paramagnetic beads. A magnet retrieves the paramagnetic beads from the solution, leaving nonhybridized single-stranded DNAs behind. Subsequently, the captured single-stranded DNA target is released from the biotinylated oligonucleotide. After release, the cDNA clone is further enriched by using a nonbiotinylated target oligonucleotide to specifically prime conversion of the single-stranded DNA. Following transformation and plating, typically 20% to 100% of the colonies represent the cDNA clone of interest. To identify

the desired cDNA clone, the colonies may be screened by colony hybridization using the ^{32}P-labeled oligonucleotide as described above for solution hybridization, or alternatively by DNA sequencing and alignment of all sequences obtained from numerous clones to determine a consensus sequence.

[00153] A nucleic acid molecule of the invention can be amplified using cDNA, mRNA, or genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to all or a portion of a nucleic acid molecule of the invention can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

[00154] In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which has a nucleotide sequence complementary to the nucleotide sequence of a nucleic acid corresponding to a marker of the invention or to the nucleotide sequence of a nucleic acid encoding a protein which corresponds to a marker of the invention. A nucleic acid molecule which is complementary to a given nucleotide sequence is one which is sufficiently complementary to the given nucleotide sequence that it can hybridize to the given nucleotide sequence thereby forming a stable duplex.

[00155] Moreover, a nucleic acid molecule of the invention can comprise only a portion of a nucleic acid sequence, wherein the full length nucleic acid sequence comprises a marker of the invention or which encodes a polypeptide corresponding to a marker of the invention. Such nucleic acids can be used, for example, as a probe or primer. The probe/primer typically is used as one or more substantially purified oligonucleotides. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 7, preferably about 15, more preferably about 25, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 or more consecutive nucleotides of a nucleic acid of the invention.

[00156] Probes based on the sequence of a nucleic acid molecule of the invention can be used to detect transcripts or genomic sequences corresponding to one or more markers of the invention. The probe comprises a label group attached thereto, *e.g.*, a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as part of a diagnostic test kit for identifying cells or tissues which mis-express the protein, such as by measuring levels of a nucleic acid molecule encoding the protein in

a sample of cells from a subject, *e.g.*, detecting mRNA levels or determining whether a gene encoding the protein has been mutated or deleted.

[00157] The invention further encompasses nucleic acid molecules that differ, due to degeneracy of the genetic code, from the nucleotide sequence of nucleic acids encoding a protein which corresponds to a marker of the invention, and thus encode the same protein.

[00158] In addition to the nucleotide sequences described herein, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequence can exist within a population (*e.g.*, the human population). Such genetic polymorphisms can exist among individuals within a population due to natural allelic variation. An allele is one of a group of genes which occur alternatively at a given genetic locus. In addition, it will be appreciated that DNA polymorphisms that affect RNA expression levels can also exist that may affect the overall expression level of that gene (*e.g.*, by affecting regulation or degradation).

[00159] As used herein, the phrase "allelic variant" refers to a nucleotide sequence which occurs at a given locus or to a polypeptide encoded by the nucleotide sequence.

[00160] As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a polypeptide corresponding to a marker of the invention. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of a given gene. Alternative alleles can be identified by sequencing the gene of interest in a number of different individuals. This can be readily carried out by using hybridization probes to identify the same genetic locus in a variety of individuals. Any and all such nucleotide variations and resulting amino acid polymorphisms or variations that are the result of natural allelic variation and that do not alter the functional activity are intended to be within the scope of the invention.

[00161] In another embodiment, an isolated nucleic acid molecule of the invention is at least 7, 15, 20, 25, 30, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 450, 550, 650, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2800, 3000, 3500, 4000, 4500, or more nucleotides in length and hybridizes under stringent conditions to a nucleic acid corresponding to a marker of the invention or to a nucleic acid encoding a protein corresponding to a marker of the invention. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 75% (80%, 85%, preferably 90%) identical to each other typically remain hybridized to each other. Such stringent conditions

are known to those skilled in the art and can be found in sections 6.3.1-6.3.6 of *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989). A preferred, non-limiting example of stringent hybridization conditions for annealing two single-stranded DNA each of which is at least about 100 bases in length and/or for annealing a single-stranded DNA and a single-stranded RNA each of which is at least about 100 bases in length, are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50-65°C. Further preferred hybridization conditions are taught in Lockhart, *et al.*, Nature Biotechnology, Volume 14, 1996 August:1675-1680; Breslauer, *et al.*, Proc. Natl. Acad. Sci. USA, Volume 83, 1986 June: 3746-3750; Van Ness, *et al.*, Nucleic Acids Research, Volume 19, No. 19, 1991 September: 5143-5151; McGraw, *et al.*, BioTechniques, Volume 8, No. 6 1990: 674-678; and Milner, *et al.*, Nature Biotechnology, Volume 15, 1997 June: 537-541, all expressly incorporated by reference.

[00162] In addition to naturally-occurring allelic variants of a nucleic acid molecule of the invention that can exist in the population, the skilled artisan will further appreciate that sequence changes can be introduced by mutation thereby leading to changes in the amino acid sequence of the encoded protein, without altering the biological activity of the protein encoded thereby. For example, one can make nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are not conserved or only semi-conserved among homologs of various species may be non-essential for activity and thus would be likely targets for alteration. Alternatively, amino acid residues that are conserved among the homologs of various species (*e.g.*, murine and human) may be essential for activity and thus would not be likely targets for alteration.

[00163] Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding a polypeptide of the invention that contain changes in amino acid residues that are not essential for activity. Such polypeptides differ in amino acid sequence from the naturally-occurring proteins which correspond to the markers of the invention, yet retain biological activity. In one embodiment, such a protein has an amino acid sequence that is at least about 40% identical, 50%, 60%, 70%, 80%, 90%, 95%, or 98% identical to the amino acid sequence of one of the proteins which correspond to the markers of the invention.

[00164] An isolated nucleic acid molecule encoding a variant protein can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of nucleic acids of the invention, such that one or more amino acid residue substitutions, additions, or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), non-polar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Alternatively, mutations can be introduced randomly along all or part of the coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for biological activity to identify mutants that retain activity. Following mutagenesis, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

[00165] The present invention encompasses antisense nucleic acid molecules, *i.e.*, molecules which are complementary to a sense nucleic acid of the invention, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule corresponding to a marker of the invention or complementary to an mRNA sequence corresponding to a marker of the invention. Accordingly, an antisense nucleic acid of the invention can hydrogen bond to (*i.e.* anneal with) a sense nucleic acid of the invention. The antisense nucleic acid can be complementary to an entire coding strand, or to only a portion thereof, *e.g.*, all or part of the protein coding region (or open reading frame). An antisense nucleic acid molecule can also be antisense to all or part of a non-coding region of the coding strand of a nucleotide sequence encoding a polypeptide of the invention. The non-coding regions ("5' and 3' untranslated regions") are the 5' and 3' sequences which flank the coding region and are not translated into amino acids.

[00166] An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides in length. An antisense nucleic acid of the

invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2- methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been sub-cloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

[00167] The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a polypeptide corresponding to a selected marker of the invention to thereby inhibit expression of the marker, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. Examples of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site or infusion of the antisense nucleic acid into an RA-associated body fluid. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered

systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

[00168] An antisense nucleic acid molecule of the invention can be an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual α-units, the strands run parallel to each other (Gaultier *et al.*, 1987, *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.*, 1987, *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987, *FEBS Lett.* 215:327-330).

[00169] The invention also encompasses ribozymes. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes as described in Haselhoff and Gerlach, 1988, *Nature* 334:585-591) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of the protein encoded by the mRNA. A ribozyme having specificity for a nucleic acid molecule encoding a polypeptide corresponding to a marker of the invention can be designed based upon the nucleotide sequence of a cDNA corresponding to the marker. For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved (see Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742). Alternatively, an mRNA encoding a polypeptide of the invention can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (see, *e.g.*, Bartel and Szostak, 1993, *Science* 261:1411-1418).

[00170] The invention also encompasses nucleic acid molecules which form triple helical structures. For example, expression of a polypeptide of the invention can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the gene encoding the polypeptide (*e.g.*, the promoter and/or enhancer) to form triple helical structures that prevent transcription of the gene in target cells. See generally Helene

(1991) *Anticancer Drug Des.* 6(6):569-84; Helene (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14(12):807-15.

[00171] In various embodiments, the nucleic acid molecules of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.*, 1996, *Bioorganic & Medicinal Chemistry* 4(1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996), *supra*; Perry-O'Keefe *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:14670-675.

[00172] PNAs can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup (1996), *supra*; or as probes or primers for DNA sequence and hybridization (Hyrup, 1996, *supra*; Perry-O'Keefe *et al.*, 1996, *Proc. Natl. Acad. Sci. USA* 93:14670-675).

[00173] In another embodiment, PNAs can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated which can combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNASE H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup, 1996, *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996), *supra,* and Finn *et al.* (1996) *Nucleic Acids Res.* 24(17):3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling

chemistry and modified nucleoside analogs. Compounds such as 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite can be used as a link between the PNA and the 5' end of DNA (Mag *et al.*, 1989, *Nucleic Acids Res.* 17:5973-88). PNA monomers are then coupled in a step-wise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.*, 1996, *Nucleic Acids Res.* 24(17):3357-63). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser *et al.*, 1975, *Bioorganic Med. Chem. Lett.* 5:1119-11124).

[00174] In other embodiments, the oligonucleotide can include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. WO 88/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (see, *e.g.*, Krol *et al.*, 1988, *Bio/Techniques* 6:958-976) or intercalating agents (see, *e.g.*, Zon, 1988, *Pharm. Res.* 5:539-549). To this end, the oligonucleotide can be conjugated to another molecule, *e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

[00175] The invention also includes molecular beacon nucleic acids having at least one region which is complementary to a nucleic acid of the invention, such that the molecular beacon is useful for quantitating the presence of the nucleic acid of the invention in a sample. A "molecular beacon" nucleic acid is a nucleic acid comprising a pair of complementary regions and having a fluorophore and a fluorescent quencher associated therewith. The fluorophore and quencher are associated with different portions of the nucleic acid in such an orientation that when the complementary regions are annealed with one another, fluorescence of the fluorophore is quenched by the quencher. When the complementary regions of the nucleic acid are not annealed with one another, fluorescence of the fluorophore is quenched to a lesser degree. Molecular beacon nucleic acids are described, for example, in U.S. Patent 5,876,930.

Recombinant Expression Vectors and Host Cells

[00176] Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a polypeptide corresponding to a marker of the invention (or a portion of such a polypeptide). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been

linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors, namely expression vectors, are capable of directing the expression of genes to which they are operably linked. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids (vectors). However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

[00177] The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell. This means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operably linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, *Methods in Enzymology: Gene Expression Technology* vol.185, Academic Press, San Diego, CA (1991). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein.

[00178] The recombinant expression vectors of the invention can be designed for expression of a polypeptide corresponding to a marker of the invention in prokaryotic (*e.g.*, *E. coli*) or eukaryotic cells (*e.g.*, insect cells {using baculovirus expression vectors}, yeast cells or mammalian cells). Suitable host cells are discussed further in Goeddel, *supra*. Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

[00179] Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988, *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

[00180] Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, 1988, *Gene* 69:301-315) and pET 11d (Studier *et al.*, p. 60-89, In *Gene Expression Technology: Methods in Enzymology* vol.185, Academic Press, San Diego, CA, 1991). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a co-expressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

[00181] One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, p. 119-128, In *Gene Expression Technology: Methods in Enzymology* vol. 185, Academic Press, San Diego, CA, 1990. Another strategy

is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, 1992, *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

[00182] In another embodiment, the expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari *et al.*, 1987, *EMBO J.* 6:229-234), pMFa (Kurjan and Herskowitz, 1982, *Cell* 30:933-943), pJRY88 (Schultz *et al.*, 1987, *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and pPicZ (Invitrogen Corp, San Diego, CA).

[00183] In another embodiment, the methods of the present invention include the generation of markers of the invention by direct chemical synthesis, rather than by production from DNA, using the protein synthetic machinery of living organisms or cell extracts containing such machinery.

[00184] Alternatively, the expression vector is a baculovirus expression vector. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith *et al.*, 1983, *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers, 1989, *Virology* 170:31-39).

[00185] In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, 1987, *Nature* 329:840) and pMT2PC (Kaufman *et al.*, 1987, *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook *et al.*, *supra.*

[00186] In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.*, 1987, *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton, 1988, *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989, *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.*, 1983, *Cell* 33:729-740; Queen and Baltimore, 1983, *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the

neurofilament promoter; Byrne and Ruddle, 1989, *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.,* 1985, *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss, 1990, *Science* 249:374-379) and the α-fetoprotein promoter (Camper and Tilghman, 1989, *Genes Dev.* 3:537-546).

[00187] The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operably linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to the mRNA encoding a polypeptide of the invention. Regulatory sequences operably linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue-specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid, or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub *et al.,* 1986, *Trends in Genetics*, Vol. 1(1).

[00188] Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

[00189] A host cell can be any prokaryotic (*e.g.*, *E. coli*) or eukaryotic cell (*e.g.*, insect cells, yeast or mammalian cells).

[00190] Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized

techniques for introducing foreign nucleic acid into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*supra*), and other laboratory manuals.

[00191] For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, for resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

[00192] A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce a polypeptide corresponding to a marker of the invention. Accordingly, the invention further provides methods for producing a polypeptide corresponding to a marker of the invention using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding a polypeptide of the invention has been introduced) in a suitable medium such that the marker is produced. In another embodiment, the method further comprises isolating the marker polypeptide from the medium or the host cell.

[00193] The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which a sequences encoding a polypeptide corresponding to a marker of the invention have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous sequences encoding a marker of the invention have been introduced into their genome or homologous recombinant animals in which endogenous gene(s) encoding a polypeptide corresponding to a marker of the invention sequences have been altered. Such animals are useful for studying the function and/or activity of the polypeptide corresponding to the marker and for identifying and/or evaluating modulators of polypeptide activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a

transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, an "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

[00194] A transgenic animal of the invention can be created by introducing a nucleic acid encoding a polypeptide corresponding to a marker of the invention into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the transgene to direct expression of the polypeptide of the invention to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, U.S. Patent No. 4,873,191 and in Hogan, *Manipulating the Mouse Embryo*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the transgene in its genome and/or expression of mRNA encoding the transgene in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying the transgene can further be bred to other transgenic animals carrying other transgenes.

[00195] To create an homologous recombinant animal, a vector is prepared which contains at least a portion of a gene encoding a polypeptide corresponding to a marker of the invention into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the gene. In a preferred embodiment, the vector is designed such that, upon homologous recombination, the endogenous gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous

recombination, the endogenous gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous protein). In the homologous recombination vector, the altered portion of the gene is flanked at its 5' and 3' ends by additional nucleic acid of the gene to allow for homologous recombination to occur between the exogenous gene carried by the vector and an endogenous gene in an embryonic stem cell. The additional flanking nucleic acid sequences are of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see, *e.g.*, Thomas and Capecchi, 1987, *Cell* 51:503 for a description of homologous recombination vectors). The vector is introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced gene has homologously recombined with the endogenous gene are selected (see, *e.g.*, Li *et al.*, 1992, *Cell* 69:915). The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see, *e.g.*, Bradley, *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, Ed., IRL, Oxford, 1987, pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Current Opinion in Bio/Technology* 2:823-829 and in PCT Publication NOS. WO 90/11354, WO 91/01140, WO 92/0968, and WO 93/04169.

[00196] In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.*, 1991, *Science* 251:1351-1355). If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a

transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

[00197] Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut *et al.* (1997) *Nature* 385:810-813 and PCT Publication NOS. WO 97/07668 and WO 97/07669.

Pharmaceutical Compositions

[00198] The nucleic acid molecules, polypeptides, and antibodies (also referred to herein as "active compounds") corresponding to a marker of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

[00199] The invention includes methods for preparing pharmaceutical compositions for modulating the expression or activity of a polypeptide or nucleic acid corresponding to a marker of the invention. Such methods comprise formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid corresponding to a marker of the invention. Such compositions can further include additional active agents. Thus, the invention further includes methods for preparing a pharmaceutical composition by formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid corresponding to a marker of the invention and one or more additional active compounds.

[00200] The invention also provides methods (also referred to herein as "screening assays") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, peptoids, small molecules or other drugs) which (a) bind to the marker, or (b) have a modulatory (*e.g.*, stimulatory or inhibitory) effect on the activity of the marker or, more specifically, (c) have a modulatory effect on the interactions of the

marker with one or more of its natural substrates (*e.g.*, peptide, protein, hormone, co-factor, or nucleic acid), or (d) have a modulatory effect on the expression of the marker. Such assays typically comprise a reaction between the marker and one or more assay components. The other components may be either the test compound itself, or a combination of test compound and a natural binding partner of the marker.

[00201] The test compounds of the present invention may be obtained from any available source, including systematic libraries of natural and/or synthetic compounds. Test compounds may also be obtained by any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive; see, *e.g.*, Zuckermann *et al.*, 1994, *J. Med. Chem.* 37:2678-85); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library and peptoid library approaches are limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, 1997, *Anticancer Drug Des.* 12:145).

[00202] Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994). *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

[00203] Libraries of compounds may be presented in solution (*e.g.*, Houghten, 1992, *Biotechniques* 13:412-421), or on beads (Lam, 1991, *Nature* 354:82-84), chips (Fodor, 1993, *Nature* 364:555-556), bacteria and/or spores, (Ladner, USP 5,223,409), plasmids (Cull *et al*, 1992, *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith, 1990, *Science* 249:386-390; Devlin, 1990, *Science* 249:404-406; Cwirla *et al,* 1990, *Proc. Natl. Acad. Sci.* 87:6378-6382; Felici, 1991, *J. Mol. Biol.* 222:301-310; Ladner, *supra.*).

[00204] In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a marker or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test

compounds which bind to a marker or biologically active portion thereof. Determining the ability of the test compound to directly bind to a marker can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to the marker can be determined by detecting the labeled marker compound in a complex. For example, compounds (*e.g.*, marker substrates) can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, assay components can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

[00205] In another embodiment, the invention provides assays for screening candidate or test compounds which modulate the activity of a marker or a biologically active portion thereof. In all likelihood, the marker can, *in vivo*, interact with one or more molecules, such as but not limited to, peptides, proteins, hormones, cofactors and nucleic acids. For the purposes of this discussion, such cellular and extracellular molecules are referred to herein as "binding partners" or marker "substrate".

[00206] One necessary embodiment of the invention in order to facilitate such screening is the use of the marker to identify its natural *in vivo* binding partners. There are many ways to accomplish this which are known to one skilled in the art. One example is the use of the marker as "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al*, 1993, *Cell* 72:223-232; Madura *et al*, 1993, *J. Biol. Chem.* 268:12046-12054; Bartel *et al* ,1993, *Biotechniques* 14:920-924; Iwabuchi *et al*, 1993 *Oncogene* 8:1693-1696; Brent WO94/10300) in order to identify other proteins which bind to or interact with the marker (binding partners) and, therefore, are possibly involved in the natural function of the marker. Such marker binding partners are also likely to be involved in the propagation of signals by the marker or downstream elements of a marker-mediated signaling pathway. Alternatively, such marker binding partners may also be found to be inhibitors of the marker.

[00207] The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that encodes a marker fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for

the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a marker-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.,* LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be readily detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the marker.

[00208] In a further embodiment, assays may be devised through the use of the invention for the purpose of identifying compounds which modulate (*e.g.,* affect either positively or negatively) interactions between a marker and its substrates and/or binding partners. Such compounds can include, but are not limited to, molecules such as antibodies, peptides, hormones, oligonucleotides, nucleic acids, and analogs thereof. Such compounds may also be obtained from any available source, including systematic libraries of natural and/or synthetic compounds. The preferred assay components for use in this embodiment is an RA marker identified herein, the known binding partner and/or substrate of same, and the test compound. Test compounds can be supplied from any source.

[00209] The basic principle of the assay systems used to identify compounds that interfere with the interaction between the marker and its binding partner involves preparing a reaction mixture containing the marker and its binding partner under conditions and for a time sufficient to allow the two products to interact and bind, thus forming a complex. In order to test an agent for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound can be initially included in the reaction mixture, or can be added at a time subsequent to the addition of the marker and its binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the marker and its binding partner is then detected. The formation of a complex in the control reaction, but less or no such formation in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the marker and its binding partner. Conversely, the formation of more complex in the presence of compound than in the control reaction indicates that the compound may enhance interaction of the marker and its binding partner.

[00210] The assay for compounds that interfere with the interaction of the marker with its binding partner may be conducted in a heterogeneous or homogeneous format.

Heterogeneous assays involve anchoring either the marker or its binding partner onto a solid phase and detecting complexes anchored to the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction between the markers and the binding partners (*e.g.*, by competition) can be identified by conducting the reaction in the presence of the test substance, *i.e.*, by adding the test substance to the reaction mixture prior to or simultaneously with the marker and its interactive binding partner. Alternatively, test compounds that disrupt preformed complexes, *e.g.*, compounds with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are briefly described below.

[00211] In a heterogeneous assay system, either the marker or its binding partner is anchored onto a solid surface or matrix, while the other corresponding non-anchored component may be labeled, either directly or indirectly. In practice, microtitre plates are often utilized for this approach. The anchored species can be immobilized by a number of methods, either non-covalent or covalent, that are typically well known to one who practices the art. Non-covalent attachment can often be accomplished simply by coating the solid surface with a solution of the marker or its binding partner and drying. Alternatively, an immobilized antibody specific for the assay component to be anchored can be used for this purpose. Such surfaces can often be prepared in advance and stored.

[00212] In related embodiments, a fusion protein can be provided which adds a domain that allows one or both of the assay components to be anchored to a matrix. For example, glutathione-S-transferase/marker fusion proteins or glutathione-S-transferase/binding partner can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and either the non-adsorbed marker or its binding partner, and the mixture incubated under conditions conducive to complex formation (*e.g.*, physiological conditions). Following incubation, the beads or microtiter plate wells are washed to remove any unbound assay components, the immobilized complex assessed either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of marker binding or activity determined using standard techniques.

[00213] Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a marker or a marker binding partner can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated marker or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). In certain embodiments, the protein-immobilized surfaces can be prepared in advance and stored.

[00214] In order to conduct the assay, the corresponding partner of the immobilized assay component is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted assay components are removed (*e.g.*, by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, can be directly labeled or indirectly labeled with, *e.g.*, a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds which modulate (inhibit or enhance) complex formation or which disrupt preformed complexes can be detected.

[00215] In an alternate embodiment of the invention, a homogeneous assay may be used. This is typically a reaction, analogous to those mentioned above, which is conducted in a liquid phase in the presence or absence of the test compound. The formed complexes are then separated from unreacted components, and the amount of complex formed is determined. As mentioned for heterogeneous assay systems, the order of addition of reactants to the liquid phase can yield information about which test compounds modulate (inhibit or enhance) complex formation and which disrupt preformed complexes.

[00216] In such a homogeneous assay, the reaction products may be separated from unreacted assay components by any of a number of standard techniques, including but not limited to: differential centrifugation, chromatography, electrophoresis and immunoprecipitation. In differential centrifugation, complexes of molecules may be separated from uncomplexed molecules through a series of centrifugal steps, due to the different sedimentation equilibria of complexes based on their different sizes and densities

(see, for example, Rivas, G., and Minton, A.P., *Trends Biochem Sci* 1993 Aug;18(8):284-7). Standard chromatographic techniques may also be utilized to separate complexed molecules from uncomplexed ones. For example, gel filtration chromatography separates molecules based on size, and through the utilization of an appropriate gel filtration resin in a column format, for example, the relatively larger complex may be separated from the relatively smaller uncomplexed components. Similarly, the relatively different charge properties of the complex as compared to the uncomplexed molecules may be exploited to differentially separate the complex from the remaining individual reactants, for example through the use of ion-exchange chromatography resins. Such resins and chromatographic techniques are well known to one skilled in the art (see, *e.g.,* Heegaard, 1998, *J Mol. Recognit.* 11:141-148; Hage and Tweed, 1997, *J. Chromatogr. B. Biomed. Sci. Appl.,* 699:499-525). Gel electrophoresis may also be employed to separate complexed molecules from unbound species (see, *e.g.,* Ausubel *et al* (eds.), In: Current Protocols in Molecular Biology, J. Wiley & Sons, New York. 1999). In this technique, protein or nucleic acid complexes are separated based on size or charge, for example. In order to maintain the binding interaction during the electrophoretic process, non-denaturing gels in the absence of reducing agent are typically preferred, but conditions appropriate to the particular interactants will be well known to one skilled in the art. Immunoprecipitation is another common technique utilized for the isolation of a protein-protein complex from solution (see, *e.g.,* Ausubel *et al* (eds.), In: Current Protocols in Molecular Biology, J. Wiley & Sons, New York. 1999). In this technique, all proteins binding to an antibody specific to one of the binding molecules are precipitated from solution by conjugating the antibody to a polymer bead that may be readily collected by centrifugation. The bound assay components are released from the beads (through a specific proteolysis event or other technique well known in the art which will not disturb the protein-protein interaction in the complex), and a second immunoprecipitation step is performed, this time utilizing antibodies specific for the correspondingly different interacting assay component. In this manner, only formed complexes should remain attached to the beads. Variations in complex formation in both the presence and the absence of a test compound can be compared, thus offering information about the ability of the compound to modulate interactions between the marker and its binding partner.

[00217] Also within the scope of the present invention are methods for direct detection of interactions between the marker and its natural binding partner and/or a test compound in a homogeneous or heterogeneous assay system without further sample manipulation.

For example, the technique of fluorescence energy transfer may be utilized (see, *e.g.*, Lakowicz *et al*, U.S. Patent No. 5,631,169; Stavrianopoulos *et al*, U.S. Patent No. 4,868,103). Generally, this technique involves the addition of a fluorophore label on a first 'donor' molecule (*e.g.*, marker or test compound) such that its emitted fluorescent energy will be absorbed by a fluorescent label on a second, 'acceptor' molecule (*e.g.*, marker or test compound), which in turn is able to fluoresce due to the absorbed energy. Alternately, the 'donor' protein molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the 'acceptor' molecule label may be differentiated from that of the 'donor'. Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, spatial relationships between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the 'acceptor' molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (*e.g.*, using a fluorimeter). A test substance which either enhances or hinders participation of one of the species in the preformed complex will result in the generation of a signal variant to that of background. In this way, test substances that modulate interactions between a marker and its binding partner can be identified in controlled assays.

[00218] In another embodiment, modulators of marker expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of mRNA or protein, corresponding to a marker in the cell, is determined. The level of expression of mRNA or protein in the presence of the candidate compound is compared to the level of expression of mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of marker expression based on this comparison. For example, when expression of marker mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of marker mRNA or protein expression. Conversely, when expression of marker mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of marker mRNA or protein expression. The level of marker mRNA or protein expression in the cells can be determined by methods described herein for detecting marker mRNA or protein.

[00219] In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a marker can be further confirmed *in vivo*, *e.g.*, in a whole animal model for cellular transformation.

[00220] This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, an marker modulating agent, an antisense marker nucleic acid molecule, an marker-specific antibody, or an marker-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

[00221] It is understood that appropriate doses of small molecule agents and protein or polypeptide agents depends upon a number of factors within the knowledge of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of these agents will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the agent to have upon the nucleic acid or polypeptide of the invention. Exemplary doses of a small molecule include milligram or microgram amounts per kilogram of subject or sample weight (*e.g.* about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram). Exemplary doses of a protein or polypeptide include gram, milligram or microgram amounts per kilogram of subject or sample weight (*e.g.* about 1 microgram per kilogram to about 5 grams per kilogram, about 100 micrograms per kilogram to about 500 milligrams per kilogram, or about 1 milligram per kilogram to about 50 milligrams per kilogram). It is furthermore understood that appropriate doses of one of these agents depend upon the potency of the agent with respect to the expression or activity to be modulated. Such appropriate doses can be determined using the assays described herein. When one or more of these agents is to be administered to an animal (*e.g.* a human) in order to modulate expression or activity of a polypeptide or

nucleic acid of the invention, a physician, veterinarian, or researcher can, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific agent employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

[00222] A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediamine-tetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampules, disposable syringes or multiple dose vials made of glass or plastic.

[00223] Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL (BASF; Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for

example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, or sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

[00224] Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a polypeptide or antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium, and then incorporating the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[00225] Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed.

[00226] Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches, and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

[00227] For administration by inhalation, the compounds are delivered in the form of an aerosol spray from a pressurized container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

[00228] Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be

permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

[00229] The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

[00230] In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes having monoclonal antibodies incorporated therein or thereon) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

[00231] It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

[00232] For antibodies, the preferred dosage is 0.1 mg/kg to 100 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg). If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is usually appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake

and tissue penetration. A method for lipidation of antibodies is described by Cruikshank *et al.* (1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193.

[00233] The nucleic acid molecules corresponding to a marker of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (U.S. Patent 5,328,470), or by stereotactic injection (see, *e.g.*, Chen *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.* retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

[00234] The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

Monitoring the Effectiveness of an Anti-RA Agent

[00235] As discussed above, the markers of the present invention can be used to assess whether RA has become refractory to an ongoing treatment (*e.g.*, a therapeutic treatment). This embodiment of the present invention relies on comparing two or more samples obtained from a patient undergoing anti-RA treatment. In general, it is preferable to obtain a first sample from the patient prior to beginning therapy and one or more samples during treatment. In such a use, a baseline of expression prior to therapy is determined and then changes in the baseline state of expression is monitored during the course of therapy. Alternatively, two or more successive samples obtained during treatment can be used without the need of a pre-treatment baseline sample. In such a use, the first sample obtained from the subject is used as a baseline for determining whether the expression of a particular gene is increasing or decreasing.

[00236] In general, when monitoring the effectiveness of a therapeutic treatment, two or more samples from the patient are examined. Preferably, three or more successively obtained samples are used, including at least one pretreatment sample.

Electronic Apparatus Readable Media and Arrays

[00237] Electronic apparatus readable media comprising a marker of the present invention is also provided. As uFsed herein, "electronic apparatus readable media" refers

to any suitable medium for storing, holding or containing data or information that can be read and accessed directly by an electronic apparatus. Such media can include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact disc; electronic storage media such as RAM, ROM, EPROM, EEPROM and the like; general hard disks and hybrids of these categories such as magnetic/optical storage media. The medium is adapted or configured for having recorded thereon a marker of the present invention.

[00238] As used herein, the term "electronic apparatus" is intended to include any suitable computing or processing apparatus or other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatus; networks, including a local area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as a personal digital assistants (PDAs), cellular phone, pager and the like; and local and distributed processing systems.

[00239] As used herein, "recorded" refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the markers of the present invention.

[00240] A variety of software programs and formats can be used to store the marker information of the present invention on the electronic apparatus readable medium. For example, the nucleic acid sequence corresponding to the markers can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like, as well as in other forms. Any number of dataprocessor structuring formats (*e.g.*, text file or database) may be employed in order to obtain or create a medium having recorded thereon the markers of the present invention.

[00241] By providing the markers of the invention in readable form, one can routinely access the marker sequence information for a variety of purposes. For example, one skilled in the art can use the nucleotide or amino acid sequences of the present invention in readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

[00242] The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has RA or a pre-disposition to RA, wherein the method comprises the steps of determining the presence or absence of a RA marker and based on the presence or absence of the RA marker, determining whether the subject has RA or a pre-disposition to RA and/or recommending a particular treatment for the RA or pre-RA condition.

[00243] The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has RA or a pre-disposition to RA associated with a RA marker wherein the method comprises the steps of determining the presence or absence of the RA marker, and based on the presence or absence of the RA marker, determining whether the subject has RA or a pre-disposition to RA, and/or recommending a particular treatment for the RA or pre-RA condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

[00244] The present invention also provides in a network, a method for determining whether a subject has RA or a pre-disposition to RA associated with a RA marker, said method comprising the steps of receiving information associated with the RA marker receiving phenotypic information associated with the subject, acquiring information from the network corresponding to the RA marker and/or RA, and based on one or more of the phenotypic information, the RA marker, and the acquired information, determining whether the subject has RA or a pre-disposition to RA. The method may further comprise the step of recommending a particular treatment for the RA or pre- RA condition

[00245] The present invention also provides a business method for determining whether a subject has RA or a pre-disposition to RA, said method comprising the steps of receiving information associated with the RA marker, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to the RA marker and/or RA, and based on one or more of the phenotypic information, the RA marker, and the acquired information, determining whether the subject has RA or a pre-disposition to RA. The method may further comprise the step of recommending a particular treatment for the RA or pre-RA condition.

[00246] The invention also includes gene and protein arrays comprising a RA marker of the present invention. The arrays can be used to assay expression of one or more genes or to assay expression of one or more proteins in the arrays. In one embodiment, the gene arrays can be used to assay gene expression in a tissue to ascertain

tissue specificity of genes in the array. In another embodiment, the protein arrays can be used to assay protein expression in a tissue to ascertain tissue specificity of proteins in the array. In this manner, several thousands of genes or proteins can be simultaneously assayed for expression. This allows a profile to be developed showing a battery of genes or proteins specifically expressed in one or more tissues.

[00247] In addition to such qualitative determination, the invention allows the quantitation of gene or protein expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes or proteins in the tissue is ascertainable. Thus, genes or proteins can be grouped on the basis of their tissue expression *per se* and level of expression in that tissue. This is useful, for example, in ascertaining the relationship of gene or protein expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene or protein expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the level of gene or protein expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

[00248] In another embodiment, the arrays can be used to monitor the time course of expression of one or more genes or proteins in the array. This can occur in various biological contexts, as disclosed herein, for example development of RA, progression of RA, and processes, such a cellular transformation associated with RA.

[00249] The arrays are also useful for ascertaining the effect of the expression of a gene or protein on the expression of other genes or proteins in the same cell or in different cells. This provides, for example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

[00250] The arrays are also useful for ascertaining differential expression patterns of one or more genes or proteins in normal and abnormal cells. This provides a battery of genes or proteins that could serve as a molecular target for diagnosis or therapeutic intervention.

Predictive Medicine

[00251] The present invention pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trails are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining the level of expression of polypeptides or nucleic acids corresponding to one or more markers of the invention, in order to determine whether an individual is at risk of developing RA. Such assays can be used for prognostic or predictive purposes to thereby prophylactically treat an individual prior to the onset of the RA.

[00252] Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs or other compounds administered either to inhibit RA or to treat or prevent any other disorder {*i.e.* in order to understand any RA progressive effects that such treatment may have}) on the expression or activity of a marker of the invention in clinical trials. These and other agents are described in further detail in the following sections.

Diagnostic Assays

[00253] An exemplary method for detecting the presence or absence of a polypeptide or nucleic acid corresponding to a marker of the invention in a biological sample involves obtaining a biological sample (*e.g.* a RA-associated body fluid) from a test subject and contacting the biological sample with a compound or an agent capable of detecting the polypeptide or nucleic acid (*e.g.*, mRNA, genomic DNA, or cDNA). The detection methods of the invention can thus be used to detect mRNA, protein, cDNA, or genomic DNA, for example, in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of a polypeptide corresponding to a marker of the invention include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence, liquid and gas chromatography, mass spectroscopy, and nuclear magnetic resonance, as well as other imaging technologies. *In vitro* techniques for detection of genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of a polypeptide corresponding to a marker of the invention include introducing into a subject a labeled antibody directed against the polypeptide. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

[00254] A general principle of such diagnostic and prognostic assays involves preparing a sample or reaction mixture that may contain a marker, and a probe, under appropriate conditions and for a time sufficient to allow the marker and probe to interact and bind, thus forming a complex that can be removed and/or detected in the reaction mixture. These assays can be conducted in a variety of ways.

[00255] For example, one method to conduct such an assay would involve anchoring the marker or probe onto a solid phase support, also referred to as a substrate, and detecting target marker/probe complexes anchored on the solid phase at the end of the reaction. In one embodiment of such a method, a sample from a subject, which is to be assayed for presence and/or concentration of marker, can be anchored onto a carrier or solid phase support. In another embodiment, the reverse situation is possible, in which the probe can be anchored to a solid phase and a sample from a subject can be allowed to react as an unanchored component of the assay.

[00256] There are many established methods for anchoring assay components to a solid phase. These include, without limitation, marker or probe molecules which are immobilized through conjugation of biotin and streptavidin. Such biotinylated assay components can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). In certain embodiments, the surfaces with immobilized assay components can be prepared in advance and stored.

[00257] Other suitable carriers or solid phase supports for such assays include any material capable of binding the class of molecule to which the marker or probe belongs. Well-known supports or carriers include, but are not limited to, glass, polystyrene, nylon, polypropylene, nylon, polyethylene, dextran, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite.

[00258] In order to conduct assays with the above mentioned approaches, the non-immobilized component is added to the solid phase upon which the second component is anchored. After the reaction is complete, uncomplexed components may be removed (*e.g.*, by washing) under conditions such that any complexes formed will remain immobilized upon the solid phase. The detection of marker/probe complexes anchored to the solid phase can be accomplished in a number of methods outlined herein.

[00259] In a preferred embodiment, the probe, when it is the unanchored assay component, can be labeled for the purpose of detection and readout of the assay, either

directly or indirectly, with detectable labels discussed herein and which are well-known to one skilled in the art.

[00260] It is also possible to directly detect marker/probe complex formation without further manipulation or labeling of either component (marker or probe), for example by utilizing the technique of fluorescence energy transfer (see, for example, Lakowicz *et al.*, U.S. Patent No. 5,631,169; Stavrianopoulos, *et al.*, U.S. Patent No. 4,868,103). A fluorophore label on the first, 'donor' molecule is selected such that, upon excitation with incident light of appropriate wavelength, its emitted fluorescent energy will be absorbed by a fluorescent label on a second 'acceptor' molecule, which in turn is able to fluoresce due to the absorbed energy. Alternately, the 'donor' protein molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the 'acceptor' molecule label may be differentiated from that of the 'donor'. Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, spatial relationships between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the 'acceptor' molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (*e.g.*, using a fluorimeter).

[00261] In another embodiment, determination of the ability of a probe to recognize a marker can be accomplished without labeling either assay component (probe or marker) by utilizing a technology such as real-time Biomolecular Interaction Analysis (BIA) (see, *e.g.*, Sjolander, S. and Urbaniczky, C., 1991, *Anal. Chem.* 63:2338-2345 and Szabo *et al.*, 1995, *Curr. Opin. Struct. Biol.* 5:699-705). As used herein, "BIA" or "surface plasmon resonance" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the mass at the binding surface (indicative of a binding event) result in alterations of the refractive index of light near the surface (the optical phenomenon of surface plasmon resonance (SPR)), resulting in a detectable signal which can be used as an indication of real-time reactions between biological molecules.

[00262] Alternatively, in another embodiment, analogous diagnostic and prognostic assays can be conducted with marker and probe as solutes in a liquid phase. In such an assay, the complexed marker and probe are separated from uncomplexed components by any of a number of standard techniques, including but not limited to: differential centrifugation, chromatography, electrophoresis and immunoprecipitation. In differential

centrifugation, marker/probe complexes may be separated from uncomplexed assay components through a series of centrifugal steps, due to the different sedimentation equilibria of complexes based on their different sizes and densities (see, for example, Rivas, G., and Minton, A.P., 1993, *Trends Biochem Sci.* 18(8):284-7). Standard chromatographic techniques may also be utilized to separate complexed molecules from uncomplexed ones. For example, gel filtration chromatography separates molecules based on size, and through the utilization of an appropriate gel filtration resin in a column format, for example, the relatively larger complex may be separated from the relatively smaller uncomplexed components. Similarly, the relatively different charge properties of the marker/probe complex as compared to the uncomplexed components may be exploited to differentiate the complex from uncomplexed components, for example through the utilization of ion-exchange chromatography resins. Such resins and chromatographic techniques are well known to one skilled in the art (see, *e.g.*, Heegaard, N.H., 1998, *J. Mol. Recognit.* Winter 11(1-6):141-8; Hage, D.S., and Tweed, S.A. *J Chromatogr B Biomed Sci Appl* 1997 Oct 10;699(1-2):499-525). Gel electrophoresis may also be employed to separate complexed assay components from unbound components (see, *e.g.*, Ausubel *et al.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1987-1999). In this technique, protein or nucleic acid complexes are separated based on size or charge, for example. In order to maintain the binding interaction during the electrophoretic process, non-denaturing gel matrix materials and conditions in the absence of reducing agent are typically preferred. Appropriate conditions to the particular assay and components thereof will be well known to one skilled in the art.

[00263] In a particular embodiment, the level of mRNA corresponding to the marker can be determined both by *in situ* and by *in vitro* formats in a biological sample using methods known in the art. The term "biological sample" is intended to include tissues, cells, biological fluids and isolates thereof, isolated from a subject, as well as tissues, cells and fluids present within a subject. Many expression detection methods use isolated RNA. For *in vitro* methods, any RNA isolation technique that does not select against the isolation of mRNA can be utilized for the purification of RNA from RA-associated body fluids (see, *e.g.*, Ausubel *et al.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York 1987-1999). Additionally, large numbers of tissue samples can readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski (1989, U.S. Patent No. 4,843,155).

[00264] The isolated mRNA can be used in hybridization or amplification assays that include, but are not limited to, Southern or Northern analyses, polymerase chain reaction analyses and probe arrays. One preferred diagnostic method for the detection of mRNA levels involves contacting the isolated mRNA with a nucleic acid molecule (probe) that can hybridize to the mRNA encoded by the gene being detected. The nucleic acid probe can be, for example, a full-length cDNA, or a portion thereof, such as an oligonucleotide of at least 7, 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to a mRNA or genomic DNA encoding a marker of the present invention. Other suitable probes for use in the diagnostic assays of the invention are described herein. Hybridization of an mRNA with the probe indicates that the marker in question is being expressed.

[00265] In one format, the mRNA is immobilized on a solid surface and contacted with a probe, for example by running the isolated mRNA on an agarose gel and transferring the mRNA from the gel to a membrane, such as nitrocellulose. In an alternative format, the probe(s) are immobilized on a solid surface and the mRNA is contacted with the probe(s), for example, in an Affymetrix gene chip array. A skilled artisan can readily adapt known mRNA detection methods for use in detecting the level of mRNA encoded by the markers of the present invention.

[00266] An alternative method for determining the level of mRNA corresponding to a marker of the present invention in a sample involves the process of nucleic acid amplification, *e.g.*, by rtPCR (the experimental embodiment set forth in Mullis, 1987, U.S. Patent No. 4,683,202), ligase chain reaction (Barany, 1991, *Proc. Natl. Acad. Sci. USA,* 88:189-193), self sustained sequence replication (Guatelli *et al.*, 1990, *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh *et al.*, 1989, *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi *et al.*, 1988, *Bio/Technology* 6:1197), rolling circle replication (Lizardi *et al.*, U.S. Patent No. 5,854,033) or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers. As used herein, amplification primers are defined as being a pair of nucleic acid molecules that can anneal to 5' or 3' regions of a gene (plus and minus strands, respectively, or vice-versa) and contain a short region in between. In general, amplification primers are from about 10 to 30 nucleotides in length and flank a region from about 50 to 200 nucleotides in length. Under appropriate

conditions and with appropriate reagents, such primers permit the amplification of a nucleic acid molecule comprising the nucleotide sequence flanked by the primers.

[00267] For *in situ* methods, mRNA does not need to be isolated from the patient sample prior to detection. In such methods, a cell or tissue sample is prepared/processed using known histological methods. The sample is then immobilized on a support, typically a glass slide, and then contacted with a probe that can hybridize to mRNA that encodes the marker.

[00268] As an alternative to making determinations based on the absolute expression level of the marker, determinations may be based on the normalized expression level of the marker. Expression levels are normalized by correcting the absolute expression level of a marker by comparing its expression to the expression of a gene that is not a marker, *e.g.*, a housekeeping gene that is constitutively expressed. Suitable genes for normalization include housekeeping genes such as the actin gene, or epithelial cell-specific genes. This normalization allows the comparison of the expression level in one sample, *e.g.*, a patient sample, to another sample, *e.g.*, a non-RA sample, or between samples from different sources.

[00269] In a method of determining the abundance of a marker in a sample compared to the normal or control, *i.e.*, to identify markers that are differentially present, the relative abundance may be determined by normalizing the signal obtained upon detecting the marker in a sample by reference to a suitable background parameter, *e.g.*, to the total protein in the sample being analyzed to an invariant marker, *i.e.*, a marker whose abundance is known to be similar in the sample being compared, or to the total signal detected from all proteins in the sample.

[00270] Alternatively, the expression level can be provided as a relative expression level. To determine a relative expression level of a marker, the level of expression of the marker is determined for 10 or more samples of normal versus RA patient sample isolates, preferably 50 or more samples, prior to the determination of the expression level for the sample in question. The mean expression level of each of the genes assayed in the larger number of samples is determined and this is used as a baseline expression level for the marker. The expression level of the marker determined for the test sample (absolute level of expression) is then divided by the mean expression value obtained for that marker. This provides a relative expression level.

[00271] Preferably, the samples used in the baseline determination will be from RA or from non-RA patient samples. The choice of the cell source is dependent on the use of

the relative expression level. Using expression found in normal tissues as a mean expression score aids in validating whether the marker assayed is RA specific (versus normal cells). In addition, as more data is accumulated, the mean expression value can be revised, providing improved relative expression values based on accumulated data. Expression data from RA patient samples provides a means for grading the severity of the RA state.

[00272] In another embodiment of the present invention, a polypeptide corresponding to a marker is detected. A preferred agent for detecting a polypeptide of the invention is an antibody capable of binding to a polypeptide corresponding to a marker of the invention, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin.

[00273] Proteins from patient samples can be isolated using techniques that are well known to those of skill in the art. The protein isolation methods employed can, for example, be such as those described in Harlow and Lane (Harlow and Lane, 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

[00274] A variety of formats can be employed to determine whether a sample contains a protein that binds to a given antibody. Examples of such formats include, but are not limited to, enzyme immunoassay (EIA), radioimmunoassay (RIA), Western blot analysis, protein arrays, antibody arrays, enzyme linked immunoabsorbant assay (ELISA), "sandwich" immunoassays, immunoprecipitation assays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays, fluorescent immunoassays and protein A immunoassays. A skilled artisan can readily adapt known protein/antibody detection methods for use in determining whether a patient sample expresses a marker of the present invention.

[00275] In one format, antibodies, or antibody fragments, can be used in methods such as Western blots, antibody arrays or immunofluorescence techniques to detect the expressed proteins. In such uses, it is generally preferable to immobilize either the antibody or proteins on a solid support. Suitable solid phase supports or carriers include any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. For protein and antibody arrays see, *e.g.* US 6,365,418, US 6,329,209, US 6,406,921, US 6,475,808 and US 6,475,809.

[00276] One skilled in the art will know many other suitable carriers for binding antibody or antigen, and will be able to adapt such support for use with the present invention. For example, protein isolated from a patient sample can be run on a polyacrylamide gel electrophoresis and immobilized onto a solid phase support such as nitrocellulose. The support can then be washed with suitable buffers followed by treatment with the detectably labeled antibody. The solid phase support can then be washed with the buffer a second time to remove unbound antibody. The amount of bound label on the solid support can then be detected by conventional means.

[00277] The invention also encompasses kits for detecting the presence of a polypeptide or nucleic acid corresponding to a marker of the invention in a biological sample (*e.g.* an RA-associated body fluid). Such kits can be used to determine if a subject is suffering from or is at increased risk of developing RA. For example, the kit can comprise a labeled compound or agent capable of detecting a polypeptide or an mRNA encoding a polypeptide corresponding to a marker of the invention in a biological sample and means for determining the amount of the polypeptide or mRNA in the sample (*e.g.*, an antibody which binds the polypeptide or an oligonucleotide probe which binds to DNA or mRNA encoding the polypeptide). Kits can also include instructions for interpreting the results obtained using the kit.

[00278] For antibody-based kits, the kit can comprise, for example: (1) a first antibody (*e.g.*, attached to a solid support) which binds to a polypeptide corresponding to a marker of the invention; and, optionally, (2) a second, different antibody which binds to either the polypeptide or the first antibody and is conjugated to a detectable label.

[00279] For oligonucleotide-based kits, the kit can comprise, for example: (1) an oligonucleotide, *e.g.*, a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence encoding a polypeptide corresponding to a marker of the invention or (2) a

pair of primers useful for amplifying a nucleic acid molecule corresponding to a marker of the invention. The kit can also comprise, *e.g.*, a buffering agent, a preservative, or a protein stabilizing agent. The kit can further comprise components necessary for detecting the detectable label (*e.g.*, an enzyme or a substrate). The kit can also contain a control sample or a series of control samples which can be assayed and compared to the test sample. Each component of the kit can be enclosed within an individual container and all of the various containers can be within a single package, along with instructions for interpreting the results of the assays performed using the kit.

Pharmacogenomics

[00280] Agents or modulators which have a stimulatory or inhibitory effect on expression of a marker of the invention can be administered to individuals to treat (prophylactically or therapeutically) RA in the patient. In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the level of expression of a marker of the invention in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

[00281] Pharmacogenomics deals with clinically significant variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, *e.g.*, Linder (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body are referred to as "altered drug action." Genetic conditions transmitted as single factors altering the way the body acts on drugs are referred to as "altered drug metabolism". These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is hemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

[00282] As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, a PM will show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

[00283] Thus, the level of expression of a marker of the invention in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a modulator of expression of a marker of the invention.

[00284] This invention also provides a process for preparing a database comprising at least one of the markers. For example, the polynucleotide sequences are stored in a digital storage medium such that a data processing system for standardized representation of the genes that identify a RA cell is compiled. The data processing system is useful to analyze gene expression between two cells by first selecting a cell suspected of being of a neoplastic phenotype or genotype and then isolating polynucleotides from the cell. The isolated polynucleotides are sequenced. The sequences from the sample are compared with the sequence(s) present in the database using homology search techniques. Greater than 90%, more preferably greater than 95% and more preferably, greater than or equal to

97% sequence identity between the test sequence and the polynucleotides of the present invention is a positive indication that the polynucleotide has been isolated from a RA cell as defined above.

[00285] In an alternative embodiment, the polynucleotides of this invention are sequenced and the information regarding sequence and in some embodiments, relative expression, is stored in any functionally relevant program, *e.g.*, in Compare Report using the SAGE software (available though Dr. Ken Kinzler at John Hopkins University). The Compare Report provides a tabulation of the polynucleotide sequences and their abundance for the samples normalized to a defined number of polynucleotides per library (say 25,000). This is then imported into MS-ACCESS either directly or via copying the data into an Excel spreadsheet first and then from there into MS-ACCESS for additional manipulations. Other programs such as SYBASE or Oracle that permit the comparison of polynucleotide numbers could be used as alternatives to MS-ACCESS. Enhancements to the software can be designed to incorporate these additional functions. These functions consist in standard Boolean, algebraic, and text search operations, applied in various combinations to reduce a large input set of polynucleotides to a manageable subset of a polynucleotide of specifically defined interest.

Monitoring Clinical Trials

[00286] Monitoring the influence of agents (*e.g.*, drug compounds) on the level of expression of a marker of the invention can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent to affect marker expression can be monitored in clinical trials of subjects receiving treatment for RA. In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of one or more selected markers of the invention in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression of the marker(s) in the post-administration samples; (v) comparing the level of expression of the marker(s) in the pre-administration sample with the level of expression of the marker(s) in the post-administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent can be

desirable to increase expression of the marker(s) to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent can be desirable to decrease expression of the marker(s) to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Surrogate Markers

[00287] The markers of the invention may serve as surrogate markers for one or more disorders or disease states or for conditions leading up to disease states, and in particular, RA. As used herein, a "surrogate marker" is an objective biochemical marker which correlates with the absence or presence of a disease or disorder, or with the progression of a disease or disorder (*e.g.*, with the presence or absence of RA symptoms). While the presence or quantity of such markers is independent of the disease, changes in the absence or presence or quantity of the marker serve as a reflection of the disease or its treatment. Therefore, these markers may serve to indicate whether a particular course of treatment is effective in lessening a disease state or disorder. Surrogate markers are of particular use when the presence or extent of a disease state or disorder is difficult to assess through standard methodologies (*e.g.*, early stage RA), or when an assessment of disease progression is desired before a potentially dangerous clinical endpoint is reached (*e.g.*, an assessment of cardiovascular disease may be made using cholesterol levels as a surrogate marker, and an analysis of HIV infection may be made using HIV RNA levels as a surrogate marker, well in advance of the undesirable clinical outcomes of myocardial infarction or fully-developed AIDS). Examples of the use of surrogate markers in the art include: Koomen *et al.* (2000) *J. Mass. Spectrom.* 35: 258-264; and James (1994) *AIDS Treatment News Archive* 209.

[00288] The markers of the invention are also useful as pharmacodynamic markers. As used herein, a "pharmacodynamic marker" is an objective biochemical marker which correlates specifically with drug effects. The presence or quantity of a pharmacodynamic marker is not related to the disease state or disorder for which the drug is being administered; therefore, the presence or quantity of the marker is indicative of the presence or activity of the drug in a subject. For example, a pharmacodynamic marker may be indicative of the concentration of the drug in a biological tissue, in that the marker is either expressed or transcribed or not expressed or transcribed in that tissue in relationship to the level of the drug. In this fashion, the distribution or uptake of the drug may be monitored by the pharmacodynamic marker. Similarly, the presence or quantity of the

pharmacodynamic marker may be related to the presence or quantity of the metabolic product of a drug, such that the presence or quantity of the marker is indicative of the relative breakdown rate of the drug *in vivo*. Pharmacodynamic markers are of particular use in increasing the sensitivity of detection of drug effects, particularly when the drug is administered in low doses. Since even a small amount of a drug may be sufficient to activate multiple rounds of marker transcription or expression, the amplified marker may be in a quantity which is more readily detectable than the drug itself. Also, the marker may be more easily detected due to the nature of the marker itself; for example, using the methods described herein, antibodies may be employed in an immune-based detection system for a protein marker, or marker-specific radiolabeled probes may be used to detect a mRNA marker. Furthermore, the use of a pharmacodynamic marker may offer mechanism-based prediction of risk due to drug treatment beyond the range of possible direct observations. Examples of the use of pharmacodynamic markers in the art include: Matsuda *et al.* US 6,033,862; Hattis *et al.* (1991) *Env. Health Perspect.* 90: 229-238; Schentag (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S21-S24; and Nicolau (1999) *Am, J. Health-Syst. Pharm.* 56 Suppl. 3: S16-S20.

[00289] The markers of the invention are also useful as pharmacogenomic markers. As used herein, a "pharmacogenomic marker" is an objective biochemical marker which correlates with a specific clinical drug response or susceptibility in a subject (see, *e.g.*, McLeod *et al.* (1999) *Eur. J. Cancer* 35(12): 1650-1652). The presence or quantity of the pharmacogenomic marker is related to the predicted response of the subject to a specific drug or class of drugs prior to administration of the drug. By assessing the presence or quantity of one or more pharmacogenomic markers in a subject, a drug therapy which is most appropriate for the subject, or which is predicted to have a greater degree of success, may be selected. For example, based on the presence or quantity of RNA or protein for specific RA markers in a subject, a drug or course of treatment may be selected that is optimized for the treatment of the specific RA likely to be present in the subject. Similarly, the presence or absence of a specific sequence mutation in marker DNA may correlate with drug response. The use of pharmacogenomic markers therefore permits the application of the most appropriate treatment for each subject without having to administer the therapy.

EXPERIMENTAL

[00290] The markers of the present invention were thus initially identified in the serum of human patients who have been diagnosed with either erosive or non-erosive RA. The markers were identified by mass spectrometry after serum samples were subjected to a series of protein depletion and fractionation steps to enrich subsets of proteins from the original serum samples. The following materials and methods describe the fundamental technologies/methodologies that were used in the marker discovery process.

Patients

[00291] Serum was collected from patients with erosive and non-erosive arthritis. Equal amounts of serum from individuals with non-erosive arthritis were pooled to create a pool of non-erosive serum for analysis. Likewise, approximately equal amounts of serum from individuals with erosive arthritis were pooled to create a pool of erosive serum for analysis. Also, equal amounts of serum from healthy individuals (ranging in age from 20-40 years old) were pooled to create a pool of healthy serum to be used as a control. This set of samples constituted a first group of serum samples for comparative analysis by mass spectrometry. Patients were sorted into erosive and non-erosive samples by the following inclusion criteria: 1) diagnosis of RA via the accepted American College of Rheumatology criteria, and 2) the age of onset of symptoms between 25 to 83. The exclusion criteria consisted of 1) a history or evidence (X-ray) of osteo arthritis, 2) systemic lupus erythematosus (SLE), 3) psoriasis or psoriatic arthritis, and 4) JRA, except in those cases with elevated rheumatoid factor.

Methods

[00292] Affinity chromatography columns was used for depletion of three abundant proteins in serum samples: a hemoglobin column for haptoglobin; protein G columns for IgG removal; and Hitrap cibacron blue columns for albumin removal. During depletion, a ConA Sepharose column is used to capture a subset of glycoproteins in serum. After depletion, samples are fractionated by size-exclusion chromatography.

Preparation of the Hemoglobin Column

[00293] Dissolve 40 mg of hemoglobin (Sigma, cat# H0267) in 1.5 mL of coupling buffer (0.2 M $NaHCO_3$, 0.5 M NaCl, pH 8.3). Desalt the solution using a HiTrap

Desalting column (Amersham Biosciences, cat# 17-1408-01) with the coupling buffer as the running buffer. Adjust the volume to the concentration of 20 mg/mL of hemoglobin.

[00294] Wash a 1 mL Hitrap NHS-activated HP column (Amersham Biosciences, cat#17-0716-01) with 5 mL of ice-cold 1 mM HCl, then immediately inject 1 mL of the hemoglobin solution and incubate at RT for 30 minutes. Wash the column with 5 mL of deactivation buffer (0.5 M ethanolamine, 0.5 M NaCl, pH 8.3) and incubate at RT for 30 minutes. Finally wash the column with 10 mL of depletion buffer (20 mM Tris/HCl, pH 7.5), and the column is ready for use. The column can be stored at 4 °C for overnight.

Preparation of the Protein L Column

[00295] Dissolve 2.5 mg of Protein L (Sigma, cat# H-3101) in 0.7 mL of coupling buffer (0.2 M NaHCO3, 0.5 M NaCl, pH 8.3). Wash one 1 mL Hitrap NHS-activated HP column (Amersham Biosciences, cat#17-0716-01) with 5 mL of ice-cold 1 mM HCl, then immediately inject 0.7 mL of the Protein L solution and incubate at RT for 30 minutes. Wash the column with 5 mL of deactivation buffer (0.5 M ethanolamine, 0.5 M NaCl, pH 8.3) and incubate at RT for 30 minutes. Finally wash the column with 10 mL of depletion buffer (20 mM Tris/HCl, pH 7.5), and the column is ready for use.

Depletion of Serum Samples

[00296] Affinity columns were prepared and set up in tandem for depletion of serum samples. In one example, columns were set up in the following manner: a hemoglobin column prepared as above; two protein G columns (1 mL each) (Amersham Biosciences, cat# 17-0404-01); a 2.5 mL ConA Sephrose (Amersham Biosciences, cat# 17-0440-03) column; and three 1 mL Cibacron Blue columns (Amersham Biosciences, cat# 17-0413-01). Wash the columns with 90 mL of depletion buffer (20 mM Tris/HCl, pH 7.5).

[00297] Dilute 1.25 mL of serum with 2.5 mL of depletion, and load the sample onto the assembled columns and wash with the depletion buffer (20 mM Tris/HCl, pH 7.5) at the flow rate of 0.5 mL/min. Flow through was collected until A_{280} returned to baseline. Freeze-dry the flow-through for 48 hours, and store the dry powder at -20 °C for the next step. ConA sepharose column was eluted with 1 M methyl mannoside in depletion buffer plus 10 mM TCEP, and effluent concentrated down to about 1 mL by using Centriplus YM-3 concentrator (Millipore, cat# 4420).

[00298] In another example, columns were set up in the following manner one hemoglobin column prepared as above; two protein G columns (1 mL each) (Amersham Biosciences, cat# 17-0404-01); one Protein L column prepared as above; one 2.5 mL ConA Sephrose (Amersham Biosciences, cat# 17-0440-03) column; and three 1 mL Cibacron Blue columns (Amersham Biosciences, cat# 17-0413-01). Wash the columns with 90 mL of depletion buffer (20 mM Tris/HCl, pH 7.5).

[00299] Dilute 1.25 mL of serum with 2.5 mL of depletion, and load the sample onto the assembled columns and wash with the depletion buffer (20 mM Tris/HCl, pH 7.5) at the flow rate of 0.5 mL/min. Collect the flow-through until A280 goes back to the baseline. Freeze-dry the flow-through for 48 hours, and store the dry powder at -20C for the next step. Elute the hemoglobin column, protein G column, protein L column, and ConA column with 0.5 M NaCl in depletion buffer until A280 returned to the baseline. Concentrate the effluent down to about 1 mL by using Centriplus YM-3 concentrator (Millipore, cat# 4420). The concentrated effluent was then combined with the freeze-dried flow-through before SEC fractionation.

SEC Fractionation of the Depleted Serum Samples

[00300] Dissolve the lyophilized flow-through sample in 1.5 mL of 6 M GdnHCl with 50 mM Tris-HCl, pH 8.0. Add 30 μL of 1 M DTT and incubate for 60 minutes at 60°C. To alkylate the sample, 150 μL of 0.5 M iodoacetamide is added. After 30 minutes of incubation in dark at RT, the alkylated sample is immediately loaded onto the SEC column. To ConA effluent, add GdnHCl solid to final concentration of 6 M, then perform reduction and alkylation as described above. Size-fractionate flow-through and effluent samples separately by using the following conditions.

[00301] The column (Superdex 200 16/60, Amersham Biosciences, cat# 17-1069-01) is pre-equilibrated with 240 mL of the running buffer (200 mM NH_4HCO_3, 8 M urea). The flowrate is 0.5 mL/min. Start collecting 5 mL fractions 76 minutes after injection. Proteins with molecular weight below 40 kDa are collected in fractions #5 to #12.

[00302] The fractions are concentrated and diluted with water to final volumes of approximately 50 μL, with final buffer composition of 50 mM NH_4HCO_3 and 2 M urea. Centriplus YM-3 and CentriconYM-3 (Millipore, cat# 4420 and 4203) are used for concentrating the fractions.

Mass spectrometry of the fractionated serum samples for discovery of protein markers

[00303] SEC fractions were digested with trypsin (2%, w/w) at 37^0C for 16 hours. The digests were desalted by using a C_{18} Vydac column. The peptide mixtures were collected and then vacuum concentrated to a final volume of ~50μL. An aliquot (10μL) of each solution was subjected to automated on-line 2D-LC/MS/MS analysis.

[00304] The 2D-LC system was composed of a capillary binary HPLC pump (Agilent), a strong cation exchange column (BioBasic SCX, 300μm x 5cm, Thermo Hypersil-Keystone), a 10-port switch valve (Valco Instruments Co.), a C_{18} desalting pre-column (150μm x 4cm) packed with Magic C_{18} material, and a C_{18} analytical column (Magic, 75μm x 15cm) . Automation between the autosampler, HPLC pump, switch valve, and mass spectrometer was accomplished by contact closure. For each SEC fraction, seven salt elution steps (10mM, 20mM, 40mM, 60mM, 80mM, 100mM, and 250mM NaCl solution containing 0.1% formic acid) were used, each elutes subsets of peptides from the SCX column. Peptides eluted from each salt step were desalted on the C_{18} pre-column, separated on C_{18} analytical column using a 3-hour gradient (4-55%B, where solvent A is 0.1% formic acid and solvent B is 90% acetonitrile with 0.1% formic acid), and subsequently analyzed by ion trap mass spectrometer with nanospray ionization.

Protein Identification

a) The raw output of mass spectra was processed using software proprietary to Millennium Pharmaceuticals Inc., called SpectrumMill. The output obtained from SpectrumMill provides an analysis of proteins present in individual SEC fractions of the original serum samples. Spectra were searched against a non-redundant NCBI mammals database. Validation of peptides was performed by either using SpectrumMill's "Automatic Validation of MS-Tag Results", by validating spectra manually or by running 1D SDS PAGE gels on serum samples.

Results

[00305] Table 1 list the markers identified using the foregoing protocol. This Table lists the markers designated with a name ("Marker"), the name the gene is commonly known by, if applicable ("Gene Name"), the data generated for each serum sample ("Erosive"; Non-Erosive"; Healthy"), the corresponding molecular weight ("protein MW (kDa)"), the corresponding GenBank Accession Number ("accession number"), the

[00306] Table 2 lists all of the markers of the invention which are over-expressed in patients with RA compared to normal individuals (*i.e.*, individuals who are not afflicted

with RA). Table 2 lists markers that are newly-associated with RA and are over-expressed in patients diagnosed with erosive or non-erosive RA. Table 2 lists preferred markers of the present invention. Table 2 lists markers which are over-expressed in serum samples of patients with RA compared to normal individuals (*i.e.*, individuals who are not afflicted with RA).

Table 1

	Erosive		Non Erosive		Healthy				
Marker #	# spectra	total intensity	# spectra	total intensity	# spectra	total intensity	mw (kDa)	Gene name	Access. no.
M1	659	7.02E+11	414	3.45E+11	587	4.61E+11	77.1	transferrin precursor PRO1557 protein	4557871
M2	488	3.86E+11	506	2.69E+11	263	5.80E+10	69.4	albumin precursor PRO0883 protein	4502027
M3	1780	3.53E+12	1615	2.73E+12	604	7.76E+11	29.0	apolipoprotein A-I precursor	178775
M4	826	1.36E+12	371	3.34E+11	762	6.53E+11	46.7	Alpha-1-antitrypsin precursor (Alpha-1 protease inhibitor) (Alpha-1-antiproteinase)	1703025
M5	242	4.80E+11	302	3.32E+11	439	5.57E+11	70.0	coagulation factor II precursor prothrombin	4503635
M6	137	1.70E+11	37	5.58E+09	155	4.01E+10	43.4	Apolipoprotein A-IV precursor (Apo-AIV)	178779
M7	496	5.42E+11	244	1.93E+11	550	5.85E+11	49.3	hemopexin	1335098
M8	21	3.98E+09	7	5.84E+08	1	1.02E+08	90.6	plasminogen	190026
M9	128	1.35E+11	86	2.60E+10	97	1.23E+10	53.0	Vitamin D-binding protein precursor (DBP) (Group-specific component) (GC-globulin) (VDB)	2119656
M10	1811	3.90E+12	270	1.38E+11	575	5.08E+11	45.2	haptoglobin	4826762
M11	106	6.20E+10	76	1.77E+10	115	8.57E+10	72.0	T-kininogen II precursor (Major acute phase protein) (Alpha-1-MAP) (Thiostatin) [Contains: T-kinin]	386852
M12	196	1.58E+11	98	4.39E+10	63	3.67E+10	63.5	I factor (complement)	1335054
M13	80	5.17E+10	90	7.13E+10	459	4.37E+11	71.0	similar to INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN H4 PRECURSOR (ITI HEAVY CHAIN H4) (INTER-ALPHA-TRYPSIN INHIBITOR FAMILY HEAVY CHAIN-RELATED PROTEIN) (IHRP) (PLASMA KALLIKREIN SENSITIVE GLYCOPROTEIN 120) (PK-120) (GP120)	7770149
M14	73	1.15E+11	46	6.15E+10	105	1.05E+11	97.7	ceruloplasmin	180249
M15	102	5.49E+10	59	1.67E+10	71	1.12E+10	70.8	alpha-2-macroglobulin	2118403
M16	51	2.43E+10	44	1.54E+10	68	3.25E+10	52.6	serine (or cysteine) proteinase inhibitor, clade C (antithrombin), member 1 antithrombin III	4502261
M17	6	1.57E	0	0.00	19	3.48E+	85.7	gelsolin (amyloidosis,	45041

		+09		E+00		09		Finnish type) Gelsolin	65
M18	64	5.61E +10	23	1.20 E+10	7	4.51E+ 08	80.2	complement component 1, r subcomponent	45024 93
M19	26	2.91E +09	29	1.48 E+09	17	1.17E+ 09	66.1	keratin 1 Keratin-1 cytokeratin 1 hair alpha protein	17318 569
M20	79	4.02E +10	23	2.58 E+09	26	8.13E+ 09	47.7	alpha-1-antichymotrypsin, precursor	14748 212
M21	330	2.51E +11	192	1.12 E+11	219	1.14E+ 11	81.3	complement component 4B preproprotein	18563 553
M22	186	4.02E +11	56	2.76 E+10	109	1.84E+ 11	34.7	Zinc-alpha-2-glycoprotein precursor (ZN-alpha-2-glycoprotein) (ZN-alpha-2-GP)	10527 4
M23	51	2.71E +10	45	1.67 E+10	52	5.64E+ 10	51.0	complement factor H related 3	21448 88
M24	83	4.76E +10	39	1.94 E+10	35	2.75E+ 10	85.6	COMPLEMENT FACTOR B PRECURSOR (C3/C5 CONVERTASE) (PROPERDIN FACTOR B) (GLYCINE-RICH BETA GLYCOPROTEIN) (GBG) (PBF2)	13278 732
M25	1206	1.65E +12	1516	2.08 E+12	1433	1.63E+ 12	12.8	TRANSTHYRETIN PRECURSOR (PREALBUMIN)	33968 5
M26	217	3.71E +11	151	1.44 E+11	183	1.68E+ 11	53.4	Ig alpha-1 chain C region	16741 036
M27	375	6.04E +11	447	9.23 E+11	330	5.34E+ 11	16.0	beta globin	45043 49
M28	77	3.27E +10	59	2.06 E+10	51	1.75E+ 10	95.0	fibrinogen, alpha chain, isoform alpha-E preproprotein fibrinogen, A alpha polypeptide	45036 89
M29	181	3.40E +11	103	5.98 E+10	88	1.19E+ 11	38.3	beta-2-glycoprotein I precursor	14771 355
M30	44	1.51E +10	12	4.27 E+09	4	7.61E+ 08	67.0	complement component 4 binding protein, alpha Complement component 4-binding protein, alpha polypeptide complement component 4-binding protein, alpha	45025 03
M31	194	2.71E +11	74	5.94 E+10	141	2.40E+ 11	48.8	clusterin (complement lysis inhibitor, SP-40,40, sulfated glycoprotein 2, testosterone-repressed prostate message 2, apolipoprotein J)	17885 5
M32	112	7.91E +10	29	1.06 E+10	20	8.22E+ 09	49.6	Ig MU chain C region	12751 4
M33	14	2.44E +09	25	4.68 E+09	24	3.64E+ 09	69.1	afamin alpha albumin	45019 87
M34	7	2.12E +09	7	1.68 E+09	24	4.86E+ 09	55.0	heparin cofactor II	13351 04
M35	62	2.07E +10	45	5.62 E+09	36	8.76E+ 09	36.2	apolipoprotein E	45573 25
M36	93	6.61E +10	74	7.15 E+10	158	1.61E+ 11	59.6	histidine-rich glycoprotein precursor histidine-proline rich glycoprotein	45044 89

M37	39	8.48E+09	11	1.24E+09	6	1.25E+09	63.0	complement component 9	179726
M38	300	1.60E+11	198	1.02E+11	205	1.19E+11	39.0	alpha-1-microglobulin/bikunin precursor Alpha-1-microglobulin/bikunin precursor inter-alpha-trypsin Alpha-1-microglobulin/bikunin precursor (inter-alpha-trypsin inhibitor, light chain protein HC)	4502067
M39	2466	5.29E+12	2054	3.04E+12	621	1.78E+12	23.5	hypothetical protein XP_092317	3721651
M40	94	5.36E+10	46	1.93E+10	66	2.55E+10	51.9	Alpha-1B-glycoprotein	112892
M41	17	1.56E+09	24	1.51E+09	7	2.73E+08	62.1	keratin 9	18587823
M42	43	3.57E+10	11	2.38E+09	64	3.12E+10	39.6	paraoxonase 1 Paraoxonase	408299
M43	296	3.55E+11	372	6.79E+11	322	3.26E+11	15.3	alpha 2 globin	4504345
M44	156	3.39E+11	94	4.99E+10	132	1.85E+11	54.3	vitronectin	14774022
M45	15	6.32E+09	18	3.38E+09	33	1.84E+10	54.6	Alpha-2-antiplasmin precursor (Alpha-2-plasmin inhibitor) (Alpha-2-PI) (Alpha-2-AP)	112907
M46	1145	3.38E+12	469	6.02E+11	617	6.72E+11	23.5	orosomucoid 1 precursor Orosomucoid-1 (alpha-1-acid glycoprotein-1) alpha-1-acid glycoprotein 1	9257232
M47	1121	2.78E+12	1233	2.24E+12	469	1.93E+12	24.7	hypothetical protein XP_092941	87890
M48	15	7.88E+09	13	4.50E+09	19	1.22E+10	62.2	peptidoglycan recognition protein L precursor	16418403
M49	6	9.05E+08	8	3.00E+09	28	1.15E+10	66.4	coagulation factor XII precursor Hageman factor	180359
M50	31	1.55E+10	7	3.12E+09	2	1.13E+09	74.2	factor H-related protein 5	180498
M51	8	1.66E+09	12	3.32E+08	3	3.29E+08	58.8	keratin 10	18588130
M52	169	6.53E+10	201	1.35E+11	5	9.74E+08	22.9	retinol-binding protein 4, plasma precursor retinol-binding protein 4, plasma retinol-binding protein 4, interstitial	5803139
M53	19	1.26E+10	1	4.44E+07	11	3.45E+09	46.3	Pigment epithelium-derived factor precursor (PEDF) (EPC-1)	12653501
M54	106	1.09E+11	53	3.23E+10	23	2.60E+10	38.1	CD5 antigen-like (scavenger receptor cysteine rich family) Spalpha	5174411
M55	25	8.66E+09	13	6.38E+09	26	1.14E+10	22.6	tetranectin (plasminogen binding protein) tetranectin (plasminogen-binding protein)	4507557
M56	6	3.30E+09	23	3.84E+09	30	6.07E+09	53.1	angiotensin precursor	2134760

M57	23	8.28E+09	7	5.78E+08	1	2.28E+08	76.7	complement component 1, s subcomponent	4502495
M58	169	1.90E+11	117	1.51E+11	74	8.00E+10	22.2	complement component 8, gamma polypeptide	4557393
M59	3	2.86E+08	1	6.73E+06	18	6.45E+09	71.4	plasma kallikrein B1 precursor Kallikrein, plasma kallikrein B plasma kallikrein 3, plasma Fletcher factor	4504877
M60	814	2.43E+12	1234	1.59E+12	1575	4.35E+12	11.2	apolipoprotein A-II precursor	4502149
M61	205	3.68E+11	183	1.75E+11	355	4.56E+11	39.3	alpha-2-HS-glycoprotein Alpha-2HS-glycoprotein	4502005
M62	144	2.21E+11	99	1.23E+11	52	5.97E+10	21.3	apolipoprotein D precursor	4502163
M63	8	4.89E+09	0	0.00E+00	17	8.50E+09	52.3	carboxypeptidase N, polypeptide 1, 50kD precursor	4503011
M64	38	1.82E+10	35	4.42E+09	30	1.02E+10	42.3	apolipoprotein L	10645201
M65	9	7.34E+08	3	1.24E+08	69	8.58E+09	94.0	similar to inter-alpha (globulin) inhibitor, H1 polypeptide	87969
M66	65	7.78E+10	5	1.96E+09	25	1.68E+10	38.2	leucine-rich alpha-2-glycoprotein	16418467
M67	27	1.13E+10	25	3.72E+09	0	0.00E+00	51.8	similar to IG GAMMA-4 CHAIN C REGION	18999465
M68	14	4.09E+09	17	3.49E+09	25	1.63E+10	71.9	fibronectin 1, isoform 2 preproprotein cold-insoluble globulin	16933544
M69	24	6.26E+09	25	9.16E+09	29	1.67E+10	52.5	coagulation factor X precursor Prothrombinase	180336
M70	288	7.33E+11	644	7.56E+11	347	5.80E+11	8.8	apolipoprotein C-III precursor	224917
M71	115	3.05E+11	83	1.45E+11	36	6.72E+10	25.4	serum amyloid P component precursor amyloid P component, serum pentaxin-related 9.5S alpha-1-glycoprotein	4502133
M72	13	1.50E+09	14	2.11E+09	31	9.23E+09	31.7	insulin-like growth factor binding protein 3	4504617
M73	208	3.90E+11	38	9.92E+09	24	5.80E+09	11.6	serum amyloid A2	13540475
M74	43	2.41E+10	39	5.41E+10	34	9.59E+09	35.4	apolipoprotein F apolipoprrotein F	4502165
M75	8	1.09E+09	6	1.54E+09	3	5.62E+08	55.2	Plasma protease C1 inhibitor precursor (C1 Inh) (C1Inh)	15029894
M76	21	1.36E+10	13	4.80E+09	18	6.78E+09	27.0	Complement factor D precursor (C3 convertase activator) (Properdin factor D) (Adipsin)	15131535
M77	0	0.00E+00	3	5.18E+07	10	1.77E+09	48.5	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 4	11437400
M78	16	3.04E+09	7	7.12E+08	3	3.89E+08	65.2	complement component 8, alpha polypeptide precursor	4557389

M79	7	4.74E+09	2	6.54E+07	8	3.72E+09	45.1	corticosteroid binding globulin precursor corticosteroid binding globulin alpha-1 antiproteinase, antitrypsin	4502595
M80	5	5.01E+09	1	2.73E+07	1	2.66E+07	62.0	complement component 8, beta polypeptide	29575
M81	14	4.45E+09	24	1.09E+10	16	4.73E+09	75.7	MASP-2 protein	5459324
M82	6	5.05E+09	7	1.01E+09	0	0.00E+00	28.9	carbonic anhydrase I carbonic dehydratase	4502517
M83	7	7.62E+08	9	2.87E+09	0	0.00E+00	21.9	peroxiredoxin 2	13631440
M84	66	5.50E+10	117	8.75E+10	95	5.72E+10	31.7	ficolin 3 precursor ficolin (collagen/fibrinogen domain-containing) 3 (Hakata antigen)	18088432
M85	62	2.71E+10	13	6.60E+09	0	0.00E+00	25.0	C-reactive protein, pentraxin-related	14728083
M86	17	4.63E+09	10	3.47E+09	28	1.41E+10	40.1	protein C (inactivator of coagulation factors Va and VIIIa)	190323
M87	150	1.19E+11	138	1.02E+11	4	1.79E+09	26.7	complement component 1, q subcomponent, beta polypeptide precursor	12722612
M88	27	2.36E+10	34	7.34E+10	9	2.19E+09	25.5	PLASMA GLUTATHIONE PEROXIDASE PRECURSOR (GSHPX-P)	121672
M89	11	7.01E+09	0	0.00E+00	5	1.89E+09	38.7	sex hormone-binding globulin Sex hormone-binding globulin (androgen binding protein)	88602
M90	21	3.45E+09	15	2.38E+09	0	0.00E+00	16.6	angiogenin, ribonuclease, RNase A family, 5 precursor Angiogenin	18307851
M91	4	5.81E+08	5	2.17E+08	0	0.00E+00	72.5	Vitamin K-dependent protein S precursor	3980130
M92	169	3.12E+11	130	1.90E+11	94	1.30E+11	18.1	similar to IMMUNOGLOBULIN J CHAIN	17446012
M93	9	2.79E+09	2	5.17E+08	3	3.76E+08	99.9	pre-alpha (globulin) inhibitor, H3 polypeptide	14735977
M94	11	2.02E+09	4	4.18E+08	5	6.34E+08	87.4	Fibronectin (FN)	182697
M95	76	1.38E+11	28	2.24E+10	9	1.87E+09	25.7	Complement C1q subcomponent, C chain precursor	399144
M96	47	1.15E+10	74	2.51E+10	63	3.23E+10	14.8	serum amyloid A4, constitutive C-SAA	10835095
M97	29	7.28E+09	15	2.69E+09	9	9.80E+08	13.2	S100 calcium-binding protein A9 calgranulin B	4506773
M98	61	2.72E+10	43	1.48E+10	40	1.13E+10	13.9		4505981
M99	54	8.01E+10	40	1.59E+10	10	1.38E+09	11.6	platelet factor 4 variant 1 Platelet factor 4, variant 1 (PF4-like)	4505735
M100	21	3.53E+09	9	4.98E+08	4	2.01E+08	10.9	S100 calcium-binding protein A8 cystic fibrosis antigen calgranulin A	4506771

M101	0	0.00E+00	0	0.00E+00	12	2.79E+09	50.7	protein Z-dependent protease inhibitor precursor protein Z-dependent protease inhibitorprecursor	7705879
M102	0	0.00E+00	4	5.50E+07	3	2.98E+07	41.8	beta actin beta cytoskeletal actin	481515
M103	128	1.46E+11	257	1.75E+11	170	1.15E+11	10.2	Apolipoprotein C-II precursor (Apo-CII)	2134777
M104	25	1.01E+10	41	2.95E+10	15	4.09E+09	13.7	beta-2-microglobulin	4757826
M105	16	1.04E+10	0	0.00E+00	2	7.69E+08	38.4	Lumican precursor (LUM) (Keratan sulfate proteoglycan)	1708878
M106	0	0.00E+00	8	3.08E+09	4	1.90E+09	26.1	soluble mannose-binding lectin precursor mannose binding protein mannose-binding lectin Mannose-binding lectin 2, soluble (opsonic defect)	14030460
M107	4	5.59E+08	6	7.65E+08	4	2.73E+08	93.5	complement component 7 precursor	4557387
M108	20	9.58E+09	14	7.61E+09	0	0.00E+00	16.5	lysozyme precursor	4557894
M109	13	1.65E+09	12	1.19E+09	8	1.76E+09	58.6	Carboxypeptidase N 83 kDa chain (Carboxypeptidase N regulatory subunit)	115877
M110	114	7.71E+10	167	6.96E+10	0	0.00E+00	26.0	complement component 1, q subcomponent, alpha polypeptide precursor complement C1q A chain precursor, complement component C1q, A chain	7705753
M111	8	7.64E+08	4	2.69E+08	0	0.00E+00	49.7	properdin P factor, complement	3183860
M112	1	1.18E+08	1	1.73E+08	13	3.54E+09	70.7	HGF activator	4504383
M113	11	1.92E+09	8	2.20E+09	3	3.36E+08	51.8	coagulation factor IX Coagulation factor IX (plasma thromboplastic component) Factor 9 Factor IX Christmas factor	4503649
M114	22	2.79E+10	5	2.72E+09	2	7.11E+08	28.4	complement component 4 binding protein, beta complement component 4-binding protein, beta Complement component 4-binding protein, beta polypeptide	4502505
M115	51	6.14E+10	48	3.51E+10	12	1.85E+10	25.2	IG KAPPA CHAIN V-III REGION NG9 PRECURSOR	3152376
M116	0	0.00E+00	0	0.00E+00	2	2.20E+08	34.6	secreted protein, acidic, cysteine-rich (osteonectin) Osteonectin (secreted protein, acidic, cysteine-rich)	4507171
M117	3	3.98E+08	0	0.00E+00	2	2.63E+07	46.3	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase,	11422666

								antitrypsin), member 7	
M118	7	1.20E+09	7	8.56E+08	0	0.00E+00	55.3	complement component C6	618466
M119	11	7.42E+08	4	3.48E+08	7	9.39E+08	58.1	prosaposin (variant Gaucher disease and variant metachromatic leukodystrophy) Prosaposin (sphingolipid activator protein-1)	11386147
M120	27	2.47E+10	9	4.39E+09	2	1.44E+09	63.5	complement component 2 precursor C3/C5 convertase	15277207
M121	7	3.57E+09	1	1.03E+08	5	5.56E+08	49.5	fibrinogen, gamma chain, isoform gamma-A precursor fibrinogen, gamma polypeptide	71827
M122	4	3.89E+08	6	2.68E+09	3	2.57E+08	17.2	17kD fetal brain protein	11641247
M123	4	1.04E+09	0	0.00E+00	1	1.61E+08	50.8	Fibrinogen beta chain precursor [Contains: Fibrinopeptide B]	223002
M124	4	5.73E+08	13	2.37E+09	10	2.02E+09	14.6	apolipoprotein C-IV	4502161
M125	5	3.81E+09	2	8.28E+07	4	9.74E+08	30.9	hypothetical protein XP_090102	18578525
M126	4	1.67E+09	0	0.00E+00	4	5.72E+08	42.1	fetuin B fetuin-like protein	7657242
M127	34	2.60E+10	9	8.23E+08	7	2.17E+09	25.3	NO_WORTHWHILE_NAMES_FOUND	284052
M128	9	3.31E+09	7	6.79E+08	12	3.17E+09	41.7	selenoprotein P precursor	11038621
M129	9	1.05E+09	10	1.48E+09	0	0.00E+00	13.8	Ribonuclease 4 precursor (RNase 4)	4506557
M130	5	6.69E+08	6	4.28E+08	0	0.00E+00	32.5	apolipoprotein B-100 precursor	178798
M131	9	5.16E+08	7	1.70E+09	8	1.82E+09	24.3	secreted phosphoprotein 2, 24kD spp24	5902118
M132	3	1.40E+08	4	4.30E+08	4	4.48E+08	20.6	apolipoprotein M	18564881
M133	1	1.28E+08	2	1.88E+08	2	1.19E+08	15.9	superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult)) Cu/Zn superoxide dismutase Superoxide dismutase-1, soluble	4507149
M134	18	1.91E+10	3	9.27E+08	10	6.97E+09	81.8	COMPLEMENT C3 ALPHA CHAIN	116596
M135	16	8.53E+09	17	2.35E+09	24	2.43E+09	7.5	INSULIN-LIKE GROWTH FACTOR II PRECURSOR (IGF-II)	1000058
M136	0	0.00E+00	1	3.27E+08	0	0.00E+00	48.4	similar to carboxypeptidase B2 (plasma)	14753775
M137	11	9.48E+09	11	7.41E+09	7	5.35E+09	12.4	hypothetical protein XP_092311	18025640
M138	16	9.42E+09	6	4.03E+09	4	4.15E+09	11.6	IG KAPPA CHAIN V-I REGION BAN	1552286
M139	16	5.29E+09	32	7.89E+09	54	1.77E+10	9.3	apolipoprotein C-I precursor	4502157

M140	8	3.68E+09	4	1.99E+09	2	7.29E+08	22.5	Neutrophil gelatinase-associated lipocalin precursor (NGAL) (P25) (25 kDa alpha-2-microglobulin-related subunit of MMP-9) (Lipocalin 2) (Oncogene 24P3)	631308
M141	1	3.68E+07	1	1.70E+07	5	4.63E+08	42.1	inter-alpha-trypsin inhibitor heavy chain	186590
M142	3	1.57E+09	2	8.06E+08	6	3.05E+09	27.6	specific granule protein (28 kDa) cysteine-rich secretory protein-3	2136189
M143	5	4.01E+09	0	0.00E+00	1	2.69E+08	40.1	CD14 antigen precursor	4557417
M144	14	1.35E+10	14	9.65E+09	1	1.15E+08	26.4	adipose most abundant gene transcript 1 adipocyte-specific secretory protein	4757760
M145	8	2.34E+09	2	7.60E+07	0	0.00E+00	10.7	small inducible cytokine subfamily A (Cys-Cys), member 14, isoform 1 precursor chemokine CC-1 chemokine CC-3	4759070
M146	11	1.02E+10	4	1.47E+09	2	1.22E+09	11.9	immunoglobulin kappa light chain variable region	4378298
M147	10	8.90E+08	6	9.05E+07	1	5.85E+06	10.2	defensin, alpha 1, preproprotein defensin 1 human neutrophil peptide 1 myeloid-related sequence	4758146
M148	1	4.85E+07	1	2.72E+07	0	0.00E+00	50.7	chromogranin A parathyroid secretory protein 1	4502805
M149	0	0.00E+00	7	1.31E+08	0	0.00E+00	68.0	hypothetical protein XP_091755	18588687
M150	14	5.23E+09	4	1.08E+09	3	2.81E+08	17.6	ribonuclease, RNase A family, 1 (pancreatic)	1360656
M151	11	6.77E+09	1	6.94E+08	3	1.68E+09	21.0	prostaglandin D2 synthase (21kD, brain)	4506251
M152	8	1.73E+09	10	7.00E+08	15	9.89E+08	7.6	insulin-like growth factor I	183120
M153	6	2.07E+09	4	9.39E+08	3	5.80E+08	24.9	similar to galectin 3 binding protein L3 antigen Mac-2-binding protein serum protein 90K	18586756
M154	0	0.00E+00	0	0.00E+00	8	6.08E+08	96.3	apolipoprotein B100	178736
M155	17	4.82E+09	8	3.31E+09	9	1.24E+09	8.8	alpha2-macroglobulin	825615
M156	7	3.54E+09	4	4.49E+09	3	1.21E+09	10.5	immunoglobulin lambda chain	422907
M157	0	0.00E+00	1	3.02E+08	0	0.00E+00	29.2	carbonic anhydrase II	4557395
M158	0	0.00E+00	1	2.58E+07	0	0.00E+00	47.2	serum deprivation response (phosphatidylserine binding protein) serum deprivation response (phosphatidylserine-binding protein)	4759082
M159	7	2.74E+09	9	9.48E+09	5	4.07E+09	11.6	immunoglobulin light chain variable region	6735444

M160	0	0.00E+00	0	0.00E+00	4	3.81E+08	66.0	insulin-like growth factor binding protein, acid labile subunit INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN COMPLEX ACID LABILE CHAIN PRECURSOR	4826772
M161	442	5.09E+11	445	2.88E+11	292	2.35E+11	25.4	TRYPSINOGEN, CATIONIC PRECURSOR (BETA-TRYPSIN)	2507249
M162	0	0.00E+00	2	6.43E+08	2	2.19E+08	10.2	hypothetical protein XP_092928	4761372
M163	6	7.55E+08	6	9.95E+08	2	1.74E+08	18.6	retinoic acid receptor responder (tazarotene induced) 2	4506427
M164	4	9.22E+08	0	0.00E+00	0	0.00E+00	77.2	similar to fibulin 1 isoform D	14779591
M165	6	1.68E+09	0	0.00E+00	0	0.00E+00	26.1	immunoglobulin kappa light chain variable region	4378342
M166	8	4.49E+09	14	1.28E+10	2	2.04E+09	28.2	hypothetical protein XP_092940	18594205
M167	1	1.15E+08	0	0.00E+00	1	6.39E+07	80.4	macrophage stimulating 1 (hepatocyte growth factor-like)	10337615
M168	3	2.44E+08	3	8.73E+07	0	0.00E+00	11.3	dermcidin precursor AIDD protein dermcidin	16751921
M169	10	3.69E+09	8	1.08E+09	2	1.96E+08	45.0	superficial zone protein	3676501
M170	4	2.52E+09	0	0.00E+00	1	1.33E+08	36.0	gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase) precursor conjugase	4503987
M171	22	9.07E+09	9	6.50E+09	6	6.63E+09	31.9	complement component C3	554423
M172	14	4.91E+09	19	5.07E+09	6	5.45E+08	15.2	IG HEAVY CHAIN V REGION HPCG13	4100372
M173	0	0.00E+00	4	1.51E+09	1	2.86E+07	62.7	hyaluronan binding protein 2 hyaluronan-binding protein hyaluronan-binding protein 2	4758502
M174	2	4.34E+07	1	1.10E+08	6	1.27E+09	10.4	SH3BGRL3-like protein	13775198
M175	3	2.90E+08	0	0.00E+00	0	0.00E+00	24.7	predicted osteoblast protein	7661714
M176	10	3.85E+09	4	1.67E+09	3	1.26E+09	9.6	IG LAMBDA CHAIN V-II REGION BUR	16075992
M177	0	0.00E+00	0	0.00E+00	2	2.85E+08	11.6	anti-c-erbB-2 immunoglobulin light chain V	1145342
M178	17	9.14E+09	4	5.44E+09	3	9.81E+08	34.0	ficolin 2 isoform a precursor ficolin (collagen/fibrinogen domain-containing lectin) 2 (hucolin) ficolin (collagen/fibrinogen domain-containing lectin) 2 hucolin	4758348
M179	2	9.79E	0	0.00	1	2.42E+	48.3	lecithin-cholesterol	38685

		+08		E+00		08		acyltransferase precursor	8
M180	0	0.00E +00	0	0.00 E+00	5	2.33E+ 08	16.9	similar to PROCESSED VARIABLE ANTIGEN	17450 519
M181	8	5.01E +09	3	1.31 E+09	2	3.75E+ 08	10.6	immunoglobulin kappa chain variable region	14268 440
M182	5	4.13E +07	0	0.00 E+00	0	0.00E+ 00	63.5	involucrin	11345 242
M183	2	5.30E +08	0	0.00 E+00	1	1.29E+ 08	52.8	disintegrin protease ADAM-like protein decysin 1	76573 19
M184	1	1.09E +08	2	5.35 E+07	0	0.00E+ 00	39.7	complement component C3	12649 541
M185	6	5.48E +09	5	5.25 E+08	0	0.00E+ 00	9.8	small inducible cytokine subfamily A (Cys-Cys), member 18, pulmonary and activation-regulated chemokine (C-C), dendritic	45068 31
M186	3	1.08E +09	0	0.00 E+00	0	0.00E+ 00	20.4	collagen XVIII	29205 35
M187	0	0.00E +00	0	0.00 E+00	6	5.45E+ 08	5.9	coagulation factor V jinjiang A2 domain	17426 605
M188	1	2.11E +07	0	0.00 E+00	3	1.82E+ 08	11.8	diazepam binding inhibitor GABA receptor modulator endozepine acyl coenzyme A binding protein	10140 853
M189	16	4.40E +10	6	5.45 E+09	0	0.00E+ 00	98.4	cleavage and polyadenylation specific factor 1, 160kD subunit	18570 089
M190	3	4.58E +09	8	1.93 E+10	2	1.13E+ 09	11.9	immunoglobulin kappa light chain variable region	43782 32
M191	2	1.98E +08	3	3.65 E+08	0	0.00E+ 00	30.6	insulin-like growth factor binding protein 5	10834 982
M192	2	1.47E +09	1	2.25 E+08	0	0.00E+ 00	12.1	immunoglobulin light chain variable region	54197 07
M193	0	0.00E +00	3	2.95 E+09	0	0.00E+ 00	12.0	immunoglobulin kappa chain V-J region	12357 65
M194	5	2.81E +08	3	8.07 E+07	1	4.13E+ 07	15.8	cystatin C (amyloid angiopathy and cerebral hemorrhage)	45031 07
M195	6	9.03E +08	9	9.87 E+08	1	1.12E+ 08	13.6	immunoglobulin variable region	16852 23
M196	5	5.50E +08	0	0.00 E+00	0	0.00E+ 00	35.4	OSTEOPONTIN PRECURSOR (BONE SIALOPROTEIN 1) (URINARY STONE PROTEIN) (SECRETED PHOSPHOPROTEIN 1) (SPP-1) (NEPHROPONTIN) (UROPONTIN)	21197 10
M197	0	0.00E +00	3	2.61 E+08	2	2.72E+ 07	13.1	immunoglobulin heavy chain variable region	31354 09
M198	6	5.50E +09	4	6.42 E+09	4	2.40E+ 09	11.7	immunoglobulin kappa chain	10660 1
M199	3	1.57E +09	0	0.00 E+00	5	1.19E+ 09	23.2	tissue inhibitor of metalloproteinase 1 precursor Erythroid-potentiating activity (tissue inhibitor of metalloproteinases)	45075 09

								erythroid potentiating activity	
M200	3	1.17E+09	6	2.47E+09	0	0.00E+00	11.5	anti-HIV gp120 antibody light chain variable region	460857
M201	1	7.33E+07	3	8.51E+07	3	5.71E+07	11.2	cystatin B (stefin B) cystatin B (liver thiol proteinase inhibitor) epilepsy, progressive myoclonic 1 (Unverricht-Lundborg type)	68783
M202	1	3.69E+08	11	1.40E+10	5	3.61E+09	12.7	immunoglobulin lambda chain variable region	587406
M203	3	3.51E+08	0	0.00E+00	0	0.00E+00	20.8	GM2 ganglioside activator protein precursor shingolipid activator protein 3 cerebroside sulfate activator protein	4504029
M204	2	1.13E+08	1	3.40E+07	0	0.00E+00	13.3	immunoglobulin heavy chain	90824
M205	1	7.76E+07	3	1.29E+08	3	6.26E+08	52.1	EGF-containing fibulin-like extracellular matrix protein 1, isoform b fibrillin-like	9665253
M206	1	1.17E+08	0	0.00E+00	2	2.81E+08	16.7	calmodulin	4885109
M207	3	1.53E+08	4	3.69E+08	0	0.00E+00	93.1	myeloperoxidase	88182
M208	0	0.00E+00	2	7.32E+08	0	0.00E+00	11.8	immunoglobulin rearranged light chain	2218124
M209	4	2.57E+08	3	1.76E+08	6	3.11E+08	17.1	CD59 antigen p18-20 (antigen identified by monoclonal antibodies 16.3A5, EJ16, EJ30, EL32 and G344)	17473237
M210	5	4.23E+09	3	1.97E+09	1	3.88E+08	11.9	immunoglobulin kappa chain V region	416338
M211	2	2.36E+07	0	0.00E+00	0	0.00E+00	9.8	IG LAMBDA CHAIN V-I REGION WAH	4324080
M212	1	5.04E+07	0	0.00E+00	2	8.53E+07	a4.9215	thymosin, beta 4	14730886
M213	8	8.90E+09	9	4.08E+09	1	5.61E+08	10.4	anti-Gd cold agglutinin monoclonal IgMK light chain variable region anti-Gd CA IgMGAS light chain variable region	545723
M214	1	3.49E+08	0	0.00E+00	0	0.00E+00	a25.0156	Thrombospondin	553801
M215	2	1.05E+09	3	3.70E+09	2	1.57E+08	86.5	coagulation factor XIII A1 subunit precursor	182837
M216	1	7.45E+07	5	2.93E+09	0	0.00E+00	12.4	hypothetical protein XP_065511	87866
M217	2	1.09E+08	2	2.41E+08	3	7.26E+08	12.9	similar to granule cell differentiation protein	17449525
M218	12	5.04E+09	3	2.74E+08	6	8.73E+08	11.4	immunoglobulin lambda light chain VJ region	3093896
M219	0	0.00E+00	1	6.65E+08	3	5.15E+08	11.6	immunoglobulin lambda-like polypeptide 1 immunoglobulin lambda-like polypeptide 1, pre-B-	871345

								cell specific immunoglobulin omega polypeptide lambda5	
M220	2	4.33E+08	1	7.52E+07	0	0.00E+00	24.0	dermatopontin precursor	14736977
M221	1	2.65E+08	2	1.42E+08	1	2.80E+08	8.3	anaphylatoxin C5a analog	1087076
M222	0	0.00E+00	2	1.96E+09	0	0.00E+00	11.9	hypothetical protein XP_092318	18092618
M223	8	6.69E+08	0	0.00E+00	4	4.31E+08	22.3	KIAA1826 protein	14042730
M224	0	0.00E+00	0	0.00E+00	7	1.93E+08	19.6	Microfibril-associated glycoprotein-2	4505089
M225	2	5.10E+09	4	2.53E+09	0	0.00E+00	11.6	immunoglobulin light chain variable region	18025666
M226	2	6.12E+08	1	9.44E+07	1	6.01E+07	11.6	immunoglobulin lambda light chain variable region	3091184
M227	0	0.00E+00	1	4.60E+08	0	0.00E+00	13.2	immunoglobulin light chain VL region	7716048
M228	1	7.89E+07	0	0.00E+00	0	0.00E+00	10.0	secretoglobin, family 1A, member 1 (uteroglobin) Uteroglobin (Clara-cell specific 10-kD protein) secretoglobin, family 1A, member 1 uteroglobin	4507809
M229	3	1.54E+09	4	3.41E+09	3	2.02E+09	12.2	Ig light chain variable domain	1864117
M230	1	2.99E+07	2	5.76E+07	0	0.00E+00	16.6	Niemann-Pick disease, type C2 Niemann-Pick disease, type C2 gene epididymal secretory protein (19.5kD)	5453678
M231	1	4.96E+06	0	0.00E+00	0	0.00E+00	11.4	immunoglobulin lambda light chain variable region	4324124
M232	3	4.86E+08	0	0.00E+00	0	0.00E+00	50.1	Ig mu chain C region	127516
M233	3	1.32E+09	0	0.00E+00	1	1.88E+08	26.4	Similar to immunoglobulin kappa constant	17511825
M234	6	3.21E+09	4	1.96E+09	3	1.76E+09	10.9	immunoglobulin light chain lambda variable region	18307308
M235	3	1.11E+09	4	2.05E+09	1	0.00E+00	11.6	immunoglobulin lambda chain variable region	987069
M236	5	2.66E+09	1	4.63E+08	0	0.00E+00	11.5	Ig kappa chain V-IV region (Dep)	106620
M237	0	0.00E+00	1	1.37E+07	2	5.92E+07	36.2	PDZ and LIM domain protein 1 (LIM domain protein CLP-36) (C-terminal LIM domain protein 1) (Elfin)	6225154
M238	5	1.83E+09	1	1.32E+08	0	0.00E+00	14.1	IG KAPPA CHAIN V-III REGION IARC/BL41 PRECURSOR	125815
M239	2	4.68E+07	1	9.71E+07	0	0.00E+00	13.6	small inducible cytokine subfamily A (Cys-Cys), member 16	7513133
M240	1	7.67E+06	1	6.71E+06	4	1.29E+08	27.9	insulin-like growth factor binding protein 4 insulin-like growth factor-binding protein 4	10835021
M241	5	2.83E	4	3.42	0	0.00E+	11.8	immunoglobulin	20722

		+09		E+09		00		rearranged light chain	74
M242	5	2.68E +09	4	1.51 E+09	6	1.07E+ 09	11.5	hypothetical protein XP_065514	43239 12
M243	5	2.91E +09	3	1.30 E+09	3	1.07E+ 09	10.6	immunoglobulin kappa light chain variable region	14625 921
M244	1	6.99E +08	3	1.40 E+09	0	0.00E+ 00	11.0	immunoglobulin kappa chain	72243 4
M245	6	2.49E +09	3	2.12 E+09	0	0.00E+ 00	9.0	immunoglobulin kappa chain variable region	15722 742
M246	3	4.55E +08	0	0.00 E+00	0	0.00E+ 00	61.3	fibrinogen A alpha polypeptide	13529 485
M247	1	2.05E +08	2	2.16 E+08	3	7.75E+ 07	12.7	immunoglobulin light chain variable region	33280 08
M248	5	8.26E +09	4	1.29 E+10	0	0.00E+ 00	9.2	NO_WORTHWHILE_NA MES_FOUND	74387 23
M249	2	3.84E +09	2	6.83 E+09	0	0.00E+ 00	12.4	immunoglobulin light chain variable region	18025 654
M250	1	1.15E +07	0	0.00 E+00	1	1.79E+ 08	37.2	aspartylglucosaminidase precursor	18333 0
M251	4	1.85E +09	0	0.00 E+00	5	2.77E+ 09	11.4	immunoglobulin kappa light chain variable region	43782 18
M252	3	2.82E +09	4	2.47 E+09	1	2.53E+ 08	10.0	immunoglobulin lambda chain	11137 128
M253	6	9.73E +08	0	0.00 E+00	1	2.27E+ 08	11.9	immunoglobulin light chain variable region	43231 83
M254	5	3.01E +09	2	1.63 E+09	2	1.01E+ 09	11.4	immunoglobulin light chain variable region	16974 735
M255	2	1.28E +09	0	0.00 E+00	0	0.00E+ 00	12.3	anti-Sm autoantibody D23K	90883
M256	2	1.20E +08	1	2.83 E+08	0	0.00E+ 00	17.6	proteoglycan 1, secretory granule Proteoglycan 1, secretory granule (platelet proteoglycan protein core)	45060 45
M257	18	3.22E +10	7	1.67 E+09	6	1.76E+ 10	40.5	hypothetical protein FLJ14497	14318 424
M258	4	7.89E +08	2	3.06 E+08	0	0.00E+ 00	9.1	immunoglobulin lambda chain variable region	16076 040
M259	2	5.51E +07	0	0.00 E+00	0	0.00E+ 00	17.2	similar to proline-rich acidic protein	16265 875
M260	0	0.00E +00	1	3.02 E+07	1	2.23E+ 07	13.9	immunoglobulin alpha-1 chain	18466 5
M261	3	3.15E +08	0	0.00 E+00	0	0.00E+ 00	10.5	hypothetical protein XP_092943	44329 76
M262	0	0.00E +00	0	0.00 E+00	2	1.51E+ 08	88.6	Unknown (protein for MGC:20375)	15779 184
M263	0	0.00E +00	1	1.91 E+08	2	5.14E+ 08	11.5	IgM light chain variable region	16735 89
M264	2	4.06E +08	1	7.17 E+07	0	0.00E+ 00	10.4	S100 A12 protein, Calgranulin C	21469 72
M265	0	0.00E +00	0	0.00 E+00	2	7.94E+ 07	56.8	glutamate carboxypeptidase-like protein 2	16751 913
M266	2	1.07E +08	0	0.00 E+00	0	0.00E+ 00	42.2	small inducible cytokine subfamily D (Cys-X3-Cys), member 1 (fractalkine, neurotactin) Small inducible cytokine subfamily D (Cys-X3-Cys), member-1	45068 57
M267	0	0.00E	2	3.53	0	0.00E+	57.5	cortactin oncogene EMS1	14250

		+00		E+07		00			668
M268	0	0.00E +00	3	1.53 E+09	0	0.00E+ 00	18.9	hypothetical protein XP_102752	18587 856
M269	0	0.00E +00	0	0.00 E+00	2	3.37E+ 08	34.0	hypothetical protein CAB56184	14249 738
M270	0	0.00E +00	2	4.00 E+08	0	0.00E+ 00	62.8	epoxide hydrolase 2, cytoplasmic	10197 684
M271	0	0.00E +00	4	4.82 E+08	2	2.24E+ 08	16.5	cystatin M Cystatin-M	45031 13
M272	0	0.00E +00	1	1.56 E+08	1	1.38E+ 08	59.6	complement component 9	90401
M273	0	0.00E +00	11	3.08 E+09	0	0.00E+ 00	63.1	5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA-2 SUBUNIT (AMPK GAMMA-2 CHAIN) (AMPK GAMMA2) (H91620P)	14285 344
M274	2	1.02E +08	2	4.12 E+08	6	1.29E+ 09	3.3	coagulation factor V	11095 907
M275	0	0.00E +00	2	3.48 E+08	0	0.00E+ 00	12.1	Ig H-V Ox1	22424 3
M276	0	0.00E +00	2	1.24 E+09	0	0.00E+ 00	30.6	T cell receptor V delta 5	67241 53
M277	0	0.00E +00	2	5.98 E+08	0	0.00E+ 00	9.5	immunoglobulin lambda chain variable region	16117 124
M278	2	4.77E +08	0	0.00 E+00	0	0.00E+ 00	42.0	chromosome 1 open reading frame 27	13097 285
M279	0	0.00E +00	0	0.00 E+00	3	1.80E+ 09	89.6	spinophilin neurabin II	16758 226
M280	2	2.13E +07	0	0.00 E+00	0	0.00E+ 00	14.3	secretory leukocyte protease inhibitor precursor seminal proteinase inhibitor mucus proteinase inhibitor antileukoproteinase	45070 65
M281	1	1.23E +08	1	2.82 E+08	0	0.00E+ 00	26.5	NO_WORTHWHILE_NA MES_FOUND	88481
M282	2	2.98E +09	0	0.00 E+00	0	0.00E+ 00	37.4	hypothetical protein FLJ20018	89230 23
M283	0	0.00E +00	4	1.48 E+09	0	0.00E+ 00	11.7	IgM light chain variable region	16735 93
M284	2	8.76E +06	0	0.00 E+00	0	0.00E+ 00	40.8	MHC class I histocompatibility antigen HLA alpha chain precursor (clone pHLA 12.4)	70075
M285	0	0.00E +00	0	0.00 E+00	2	1.68E+ 09	34.6	follistatin-related protein precursor	13242 265
M286	4	2.39E +09	4	2.49 E+09	1	6.02E+ 07	11.4	immunoglobulin light chain variable region	18025 716
M287	0	0.00E +00	2	4.16 E+08	0	0.00E+ 00	10.2	immunoglobulin lambda light chain variable region	18041 862
M288	2	7.67E +08	0	0.00 E+00	0	0.00E+ 00	42.8	hypothetical protein XP_091589	18587 574
M289	2	2.17E +07	0	0.00 E+00	0	0.00E+ 00	18.2	NO_WORTHWHILE_NA MES_FOUND	13358 946
M290	1	4.74E +07	0	0.00 E+00	0	0.00E+ 00	62.0	apolipoprotein B100 precursor	17879 2
M291	0	0.00E +00	3	3.86 E+10	0	0.00E+ 00	70.4	NO_WORTHWHILE_NA MES_FOUND	12850 167

M292	0	0.00E+00	6	5.43E+08	0	0.00E+00	26.8	similar to basic helix-loop-helix domain containing, class B5 BETA3 protein basic helix-loop-helix (bHLH) gene, class B, Beta3	17466797
M293	0	0.00E+00	1	3.33E+08	0	0.00E+00	10.0	IG KAPPA CHAIN V-I REGION AU	6578182
M294	4	6.30E+09	0	0.00E+00	0	0.00E+00	98.5	lipin 1	17444449
M295	1	1.71E+08	1	1.41E+08	1	3.75E+08	11.8	immunoglobulin kappa chain variable region	5578814
M296	0	0.00E+00	1	5.35E+08	0	0.00E+00	16.4	similar to Unknown (protein for IMAGE:2905327)	182516
M297	0	0.00E+00	0	0.00E+00	1	2.20E+08	28.7	hypothetical protein MGC3279 similar to collectins	13128972
M298	0	0.00E+00	0	0.00E+00	2	1.63E+08	51.5	Muscarinic acetylcholine receptor M2	14194439
M299	1	0.00E+00	0	0.00E+00	1	2.49E+08	11.8	anti-cardiolipin immunoglobulin light chain	18092608
M300	1	1.50E+08	1	3.50E+07	0	0.00E+00	10.0	Stromal cell-derived factor 1 precursor (SDF-1) (Pre-B cell growth stimulating factor) (PBSF)	10834988
M301	0	0.00E+00	1	1.67E+08	1	4.12E+08	10.6	immunoglobulin kappa light chain variable region	4378356
M302	0	0.00E+00	1	2.61E+08	0	0.00E+00	11.5	immunoglobulin light chain V-J region	1944483
M303	0	0.00E+00	8	2.77E+09	1	1.84E+08	9.5	immunoglobulin kappa chain	11137021
M304	1	1.28E+08	0	0.00E+00	0	0.00E+00	83.0	cartilage oligomeric matrix protein presursor	14766987
M305	0	0.00E+00	1	1.09E+08	0	0.00E+00	5.7	serine protease inhibitor, Kazal type 1	224571
M306	1	2.96E+08	0	0.00E+00	0	0.00E+00	15.6	immunoglobulin heavy chain variable region	5679468
M307	0	0.00E+00	0	0.00E+00	1	5.83E+07	84.4	calcium channel, voltage-dependent, alpha 2/delta subunit 1	4454526
M308	0	0.00E+00	1	1.17E+08	0	0.00E+00	12.5	similar to alpha-1,3(6)-mannosylglycoprotein beta-1,6-N-acetyl-glucosaminyltransferase Mannosyl (alpha-1,6-)-glycoprotein beta-1,6-N-acetyl- alpha-mannoside beta-1,6-N-acetylglucosaminyltransferase	14751814
M309	1	1.15E+09	0	0.00E+00	0	0.00E+00	11.7	immunoglobulin light chain V-region	6648588
M310	1	9.22E+07	1	5.65E+08	1	3.79E+08	11.9	IG KAPPA CHAIN V-I REGION MEV	125776
M311	3	3.32E+08	1	1.08E+09	0	0.00E+00	10.3	anticardiolipin immunoglobulin light chain	11118905
M312	0	0.00E	0	0.00	1	1.79E+	9.0	immunoglobulin kappa	16076

		+00		E+00		08		chain variable region	161
M313	1	2.38E +08	2	1.81 E+09	3	2.20E+ 09	11.5	immunoglobulin lambda chain variable region	99683 88
M314	0	0.00E +00	0	0.00 E+00	1	2.48E+ 07	23.8	ephrin A1 precursor eph-related receptor tyrosine kinase ligand 1 (tumor necrosis factor, alpha-induced protein 4)	47582 46
M315	0	0.00E +00	1	1.21 E+08	1	9.19E+ 07	11.3	immunoglobulin V lambda/J lambda light chain	66437 93
M316	0	0.00E +00	0	0.00 E+00	2	3.00E+ 08	9.9	immunoglobulin lambda chain variable region	96633 09
M317	3	1.27E +09	0	0.00 E+00	0	0.00E+ 00	16.0	immunoglobulin kappa chain	57660 0
M318	1	2.65E +08	1	1.48 E+08	0	0.00E+ 00	42.2	surfactant, pulmonary-associated protein B Pulmonary surfactant-associated protein B, 18kD	71980
M319	0	0.00E +00	1	9.70 E+08	0	0.00E+ 00	12.1	IgG kappa light chain variable region	23068 93
M320	0	0.00E +00	1	6.44 E+08	0	0.00E+ 00	10.3	immunoglobulin lambda chain variable region	49736 3
M321	0	0.00E +00	1	3.25 E+07	0	0.00E+ 00	14.7	beta-galactosidase binding lectin precursor Lectin, galactose-binding, soluble, 1 galectin	45049 81
M322	0	0.00E +00	1	2.03 E+07	0	0.00E+ 00	11.9	FK506 binding protein 2 (13 kDa)	17985 953
M323	3	8.25E +08	2	4.62 E+08	1	1.67E+ 08	26.1	Trypsin I precursor (Cationic trypsinogen)	45061 45
M324	0	0.00E +00	0	0.00 E+00	3	2.35E+ 09	10.7	immunoglobulin kappa chain	33688
M325	5	3.68E +09	1	1.07 E+09	2	8.27E+ 08	11.1	immunoglobulin kappa light chain variable region	43239 36
M326	3	2.17E +08	1	1.31 E+07	0	0.00E+ 00	12.6	IG HEAVY CHAIN V-III REGION JON	12385 9
M327	0	0.00E +00	5	3.24 E+09	1	1.04E+ 08	11.3	immunoglobulin lambda light chain variable region	43242 68
M328	5	6.01E +09	0	0.00 E+00	2	1.10E+ 09	11.7	immunogloblin light chain	19057 99
M329	0	0.00E +00	0	0.00 E+00	1	4.03E+ 08	11.9	immunoglobulin variable region, kappa light chain	25979 36
M330	0	0.00E +00	1	5.08 E+07	0	0.00E+ 00	11.2	ARS component B precursor anti-neoplastic urinary protein secreted Ly6/uPAR related protein 1 ARS(component B)-81/S	99669 07
M331	1	0.00E +00	3	2.86 E+09	0	0.00E+ 00	13.7	amyloid lambda light chain variable region	41036 51
M332	5	3.37E +09	5	2.89 E+09	1	5.61E+ 08	12.2	IG KAPPA CHAIN V-I REGION NI	12577 7
M333	2	5.71E +08	0	0.00 E+00	0	0.00E+ 00	11.8	anti-DNA immunoglobulin light chain	17303 05
M334	0	0.00E +00	1	4.15 E+07	0	0.00E+ 00	44.1	secreted and transmembrane 1 precusor	13654 162
M335	1	2.57E +07	0	0.00 E+00	0	0.00E+ 00	11.7	hypothetical protein XP_058875	18604 158
M336	3	1.47E +09	0	0.00 E+00	1	1.81E+ 08	12.3	immunoglobulin kappa light chain variable region	57312 29

								B3	
M337	3	2.58E+09	0	0.00E+00	0	0.00E+00	11.9	immunoglobulin light chain variable region	18025576
M338	0	0.00E+00	0	0.00E+00	1	3.87E+07	14.1	immunoglobulin lambda-chain	434040
M339	1	5.22E+07	1	7.28E+06	1	2.94E+07	12.1	immunoglobulin lambda light chain variable region	3388063
M340	0	0.00E+00	1	1.82E+08	1	5.54E+07	10.0	immunoglobulin lambda chain	11137136
M341	1	2.95E+08	1	1.38E+08	1	4.91E+08	12.4	immunoglobulin lambda-chain subgroup II	1321596
M342	0	0.00E+00	1	4.17E+08	0	0.00E+00	9.3	immunoglobulin lambda chain variable region	16075968
M343	1	7.87E+08	0	0.00E+00	0	0.00E+00	17.4	dJ581P3.2 (attractin (with dipeptidylpeptidase IV activity))	7711012
M344	0	0.00E+00	4	1.32E+09	2	1.75E+08	11.5	IG LAMBDA CHAIN V-I REGION NEW	126541
M345	1	7.87E+08	0	0.00E+00	0	0.00E+00	12.0	immunoglobulin light chain variable region	19744548
M346	0	0.00E+00	0	0.00E+00	2	8.78E+07	11.4	immunoglobulin lambda light chain variable region	6643493
M347	1	3.48E+07	1	3.37E+07	0	0.00E+00	11.3	S100 calcium binding protein A14 (calgizzarin) S100 calcium-binding protein A14 (calgizzarin)	10567826
M348	0	0.00E+00	1	2.40E+07	2	2.76E+08	21.6	hypothetical protein XP_091231	18598452
M349	1	3.26E+08	1	6.19E+07	0	0.00E+00	12.3	anti-human chorionic gonadotropin monoclonal antibody AB4 Ig kappa light chain variable region	3493267
M350	1	1.11E+08	0	0.00E+00	0	0.00E+00	52.9	lipopolysaccharide binding protein lipopolysaccharide-binding protein	18490598
M351	3	1.05E+09	7	4.95E+09	2	1.13E+09	11.7	hepatitis B surface antigen antibody	183959
M352	2	1.64E+09	1	5.70E+08	0	0.00E+00	10.7	immunoglobulin light chain variable region	3153381
M353	0	0.00E+00	1	9.74E+07	2	1.01E+08	11.7	thioredoxin	4507745
M354	3	3.94E+08	1	3.92E+08	2	7.06E+08	11.4	immunoglobulin kappa light chain variable region	9246545
M355	0	0.00E+00	6	3.67E+09	0	0.00E+00	10.3	immunoglobulin light chain variable region	3927988
M356	0	0.00E+00	2	5.89E+09	0	0.00E+00	2.5	NO_WORTHWHILE_NAMES_FOUND	106612
M357	0	0.00E+00	0	0.00E+00	1	1.97E+08	33.1	asialoglycoprotein receptor 2 isoform c asialoglycoprotein receptor H2 hepatic lectin H2	18426877
M358	1	1.38E+08	0	0.00E+00	1	6.57E+08	10.9	immunoglobulin lambda chain precursor	575243
M359	0	0.00E+00	2	5.29E+08	1	2.59E+08	9.0	immunoglobulin kappa light chain variable region	18041674
M360	0	0.00E+00	1	1.48E+09	0	0.00E+00	11.9	immunoglobulin light chain variable region	8777889
M361	4	2.66E+09	2	7.79E+08	2	2.75E+08	11.9	immunoglobulin kappa, VJ region	1322200

M362	0	0.00E+00	1	4.28E+08	0	0.00E+00	11.5	immunoglobulin lambda light chain variable region	6643601
M363	0	0.00E+00	0	0.00E+00	1	5.61E+08	11.2	immunoglobulin kappa chain variable region	12655666
M364	2	3.36E+09	2	1.80E+09	0	0.00E+00	11.5	immunoglobulin kappa chain	722428
M365	0	0.00E+00	1	7.37E+08	0	0.00E+00	11.5	immunoglobulin kappa light chain variable region	4323908
M366	0	0.00E+00	4	3.54E+09	0	0.00E+00	11.6	This CDS feature is included to show the translation of the corresponding V_region. Presently translation qualifiers on V_region features are illegal.	681900
M367	2	3.67E+09	0	0.00E+00	1	3.24E+08	10.8	immunoglobulin kappa chain	722526
M368	0	0.00E+00	1	2.62E+08	0	0.00E+00	11.9	immunoglobulin lambda light chain variable region	3142596
M369	0	0.00E+00	0	0.00E+00	1	3.60E+08	11.9	immunoglobulin kappa, VJ region	1322204
M370	0	0.00E+00	1	1.30E+08	0	0.00E+00	11.4	immunoglobulin lambda light chain variable region	3091164
M371	0	0.00E+00	1	4.96E+07	0	0.00E+00	11.8	platelet-derived growth factor alpha, isoform 2, preproprotein	6119621
M372	0	0.00E+00	0	0.00E+00	1	1.33E+08	9.4	immunoglobulin kappa light chain variable region	18041788
M373	1	6.15E+07	0	0.00E+00	0	0.00E+00	18.0	IMMUNOGLOBULIN J CHAIN PRECURSOR	13543748
M374	0	0.00E+00	0	0.00E+00	1	6.20E+07	11.8	immunoglobulin light chain variable region	18698393
M375	0	0.00E+00	0	0.00E+00	1	1.87E+07	79.2	C4/C2 activating component of Ra-reactive factor	14735142
M376	0	0.00E+00	0	0.00E+00	1	1.51E+07	28.5	tropomyosin 4	4507651
M377	3	1.44E+09	1	3.16E+08	0	0.00E+00	10.2	immunoglobulin lambda chain variable region	497338
M378	1	3.92E+07	0	0.00E+00	0	0.00E+00	36.4	Transaldolase	17511894
M379	0	0.00E+00	1	5.13E+07	0	0.00E+00	25.4	hypothetical protein XP_088290	18088480
M380	0	0.00E+00	0	0.00E+00	1	1.96E+08	11.5	anti-DNA immunoglobulin light chain IgG	1870412
M381	2	2.51E+08	2	1.43E+09	1	3.20E+08	9.5	immunoglobulin lambda light chain variable region	9714348
M382	2	4.64E+08	2	8.59E+08	2	9.80E+08	11.6	anti-c-erbB-2 immunoglobulin light chain V region	1145216
M383	1	7.54E+07	0	0.00E+00	0	0.00E+00	11.8	immunoglobulin heavy chain variable region	17384988
M384	1	4.25E+07	0	0.00E+00	0	0.00E+00	12.7	immunoglobulin light chain variable region	3328006
M385	1	6.97E+08	0	0.00E+00	0	0.00E+00	10.6	immunoglobulin kappa chain	722422
M386	1	9.43E+08	1	1.09E+08	1	4.28E+08	11.7	immunoglobulin kappa light chain variable region	4323960
M387	0	0.00E+00	0	0.00E+00	1	1.43E+08	10.3	immunoglobulin kappa chain variable region	12655720

M388	0	0.00E+00	1	6.54E+08	0	0.00E+00	12.3	immunoglobulin kappa chain variable region	5578792
M389	3	3.05E+09	0	0.00E+00	0	0.00E+00	11.6	immunoglobulin lambda light chain variable region	3091192
M390	2	2.39E+09	1	3.19E+08	0	0.00E+00	8.8	immunoglobulin lambda chain variable region	16076070
M391	1	6.13E+08	1	8.21E+08	1	1.50E+08	14.5	Ig kappa light chain (VJC)	441357
M392	3	3.07E+09	1	1.06E+08	0	0.00E+00	11.6	immunoglobulin kappa light chain variable region	9246636
M393	2	1.02E+09	0	0.00E+00	0	0.00E+00	47.7	chromosome 20 open reading frame 3	9836652
M394	0	0.00E+00	0	0.00E+00	1	7.16E+08	10.6	immunoglobulin V lambda/J lambda light chain	6643721
M395	2	3.32E+08	0	0.00E+00	0	0.00E+00	12.3	immunoglobulin light chain variable region	19744484
M396	0	0.00E+00	0	0.00E+00	2	7.66E+07	10.8	immunoglobulin lambda light chain variable region	4324288
M397	0	0.00E+00	0	0.00E+00	1	2.61E+08	11.7	Ig lambda-chain V-region	1335383
M398	0	0.00E+00	1	4.27E+08	0	0.00E+00	11.6	anti-c-erbB-2 immunoglobulin light chain V	1145332
M399	1	1.89E+08	0	0.00E+00	0	0.00E+00	11.4	immunoglobulin V lambda/J lambda light chain	6643883
M400	1	9.70E+07	0	0.00E+00	0	0.00E+00	22.0	CA11	9665240
M401	1	0.00E+00	1	3.20E+08	0	0.00E+00	9.6	immunoglobulin lambda chain variable region	15722929
M402	0	0.00E+00	1	1.47E+09	0	0.00E+00	8.6	immunoglobulin kappa chain variable region	15722831
M403	0	0.00E+00	1	5.87E+08	0	0.00E+00	12.1	Ig light chain V region	481497
M404	1	1.53E+09	0	0.00E+00	0	0.00E+00	11.6	immunoglobulin lambda light chain variable region	3142452
M405	1	1.23E+07	0	0.00E+00	0	0.00E+00	8.1	WAP four-disulfide core domain 2, isoform 5 epididymis-specific, whey-acidic protein type, four-disulfide core WAP domain containing protein HE4-V4 major epididymis-specific protein E4 epididymal secretory protein E4	18379360
M406	1	2.88E+08	3	4.06E+09	0	0.00E+00	1.5	hemoglobin beta chain beta-globin	239718
M407	1	5.12E+06	0	0.00E+00	0	0.00E+00	17.1	Myoglobin	229361
M408	0	0.00E+00	1	2.11E+08	0	0.00E+00	9.1	immunoglobulin kappa chain variable region	16116925
M409	3	2.57E+09	0	0.00E+00	0	0.00E+00	11.2	immunoglobulin lambda chain variable region	12655763
M410	2	1.25E+09	0	0.00E+00	0	0.00E+00	12.1	IG KAPPA CHAIN V-II REGION MIL	125786
M411	1	0.00E+00	1	3.28E+08	0	0.00E+00	9.3	immunoglobulin lambda chain variable region	16075940
M412	0	0.00E	0	0.00	1	6.15E+	15.2	immunoglobulin light	18593

		+00		E+00		08		chain variable region	5
M413	1	5.16E +08	1	3.67 E+08	1	1.75E+ 08	11.6	immunoglobulin kappa light chain variable region	43238 94
M414	1	8.47E +07	0	0.00 E+00	0	0.00E+ 00	24.5	Alpha-S1 casein precursor	21193 98
M415	1	6.06E +08	0	0.00 E+00	0	0.00E+ 00	9.9	immunoglobulin kappa light chain V region	42618 43
M416	1	3.91E +08	0	0.00 E+00	0	0.00E+ 00	10.4	similar to V4-1	10945 908
M417	2	6.54E +08	3	7.89 E+08	3	1.65E+ 08	39.7	Heterogeneous nuclear ribonucleoprotein A3 (hnRNP A3) (D10S102)	17106 27
M418	1	2.08E +08	0	0.00 E+00	2	7.14E+ 08	8.7	immunoglobulin kappa chain variable region	10637 159
M419	1	1.03E +08	0	0.00 E+00	1	1.86E+ 08	11.7	immunoglobulin lambda light chain variable region	43242 24
M420	3	9.85E +08	7	8.04 E+08	5	2.93E+ 09	26.7	hypothetical protein XP_105996	18557 302
M421	1	8.54E +08	0	0.00 E+00	0	0.00E+ 00	11.5	immunoglobulin kappa light chain variable region	43238 46
M422	3	8.25E +08	1	3.11 E+08	0	0.00E+ 00	9.7	immunoglobulin lambda chain	28654 78
M423	9	2.93E +09	0	0.00 E+00	1	6.18E+ 07	22.7	similar to mouse Glt3 or D. malanogaster transcription factor IIB	83928 75
M424	0	0.00E +00	0	0.00 E+00	2	5.06E+ 08	14.3	immunoglobulin gamma chain (BAB4-L)	39281 82
M425	1	1.73E +09	0	0.00 E+00	0	0.00E+ 00	13.2	Ig G VL JEL44, anti-sugar phosphotransferase	22779 3
M426	1	2.81E +07	0	0.00 E+00	0	0.00E+ 00	11.8	immunoglobulin kappa-chain VK-1	19765 2
M427	1	1.22E +09	0	0.00 E+00	0	0.00E+ 00	11.5	immunoglobulin kappa chain	18604 8
M428	1	1.06E +08	0	0.00 E+00	0	0.00E+ 00	75.4	coagulation factor XIII, beta subunit	17941 7
M429	0	0.00E +00	0	0.00 E+00	1	2.59E+ 08	9.6	immunoglobulin kappa chain	10655 5
M430	2	1.44E +09	0	0.00 E+00	0	0.00E+ 00	9.8	immunoglobulin lambda chain	27919 55
M431	0	0.00E +00	3	2.24 E+09	1	5.50E+ 08	11.9	immunoglobulin light chain variable region	18698 419
M432	0	0.00E +00	1	3.62 E+08	0	0.00E+ 00	11.2	immunoglobulin light chain variable region	54196 91
M433	0	0.00E +00	1	7.04 E+07	0	0.00E+ 00	11.7	immunoglobulin light chain variable region	54196 95
M434	0	0.00E +00	0	0.00 E+00	1	3.54E+ 07	30.0	LIM and SH3 protein 1	67545 08
M435	3	6.52E +08	1	2.97 E+08	2	5.94E+ 08	14.1	This CDS feature is included to show the translation of the corresponding V_region. Presently translation qualifiers on V_region features are illegal	79079 5
M436	0	0.00E +00	1	4.96 E+08	0	0.00E+ 00	8.1	immunoglobulin lambda chain variable region	16075 614
M437	0	0.00E +00	1	5.88 E+07	0	0.00E+ 00	18.5	LEPTIN (OBESITY FACTOR)	21355 55
M438	0	0.00E +00	0	0.00 E+00	1	2.43E+ 08	13.5	Ig kappa light chain (VJC)	44139 5

M439	1	1.17E+08	0	0.00E+00	1	4.45E+08	9.8	Ig kappa chain VKIII-JK5	470514
M440	0	0.00E+00	0	0.00E+00	1	2.08E+08	10.6	NO_WORTHWHILE_NAMES_FOUND	625508
M441	0	0.00E+00	0	0.00E+00	1	2.45E+07	16.0	Hemoglobin beta chain	122572
M442	0	0.00E+00	1	3.13E+08	0	0.00E+00	11.9	immunoglobulin kappa chain	434696
M443	3	1.83E+09	0	0.00E+00	2	1.45E+09	6.6	NO_WORTHWHILE_NAMES_FOUND	106589
M444	0	0.00E+00	2	4.12E+08	0	0.00E+00	9.6	immunoglobulin lambda chain variable region	16075944
M445	0	0.00E+00	1	2.91E+08	0	0.00E+00	10.2	immunoglobulin lambda light chain variable region	4337080
M446	1	1.99E+08	2	1.04E+09	2	1.03E+09	31.5	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial precursor (Isocitric dehydrogenase) (NAD+-specific ICDH)	5031777
M447	0	0.00E+00	0	0.00E+00	12	3.84E+09	26.9	Unknown (protein for MGC:27742)	17391195
M448	2	1.50E+09	0	0.00E+00	0	0.00E+00	9.1	immunoglobulin lambda chain variable region	9663311
M449	0	0.00E+00	0	0.00E+00	1	7.32E+07	8.6	immunoglobulin lambda chain variable region	16076026
M450	1	3.51E+09	0	0.00E+00	0	0.00E+00	10.1	immunoglobulin light chain	3123582
M451	1	4.66E+08	0	0.00E+00	0	0.00E+00	35.0	Fc fragment of IgG, low affinity IIa, receptor for (CD32)	31336
M452	1	5.87E+06	0	0.00E+00	0	0.00E+00	64.2	similar to non-specific cross reacting antigen	17456384
M453	7	1.19E+09	3	4.58E+07	2	9.53E+07	16.6	IgA heavy chain variable region	13347047
M454	0	0.00E+00	1	2.34E+08	0	0.00E+00	50.7	TYROSINE-PROTEIN KINASE CSK (C-SRC KINASE)	417209
M455	0	0.00E+00	4	6.19E+07	0	0.00E+00	23.9	hairy/enhancer of split 6	14009498
M456	0	0.00E+00	0	0.00E+00	1	3.23E+08	8.1	immunoglobulin kappa chain	722548
M457	1	8.14E+08	0	0.00E+00	1	2.19E+08	13.7	IgA1 kappa light chain	6110570
M458	2	5.96E+08	0	0.00E+00	0	0.00E+00	11.3	IG LAMBDA CHAIN V-V REGION DEL	126571
M459	0	0.00E+00	0	0.00E+00	1	1.63E+08	11.6	immunoglobulin kappa light chain variable region	4378330
M460	2	3.88E+08	0	0.00E+00	0	0.00E+00	10.1	immunoglobulin lambda light chain variable region	4337017
M461	0	0.00E+00	0	0.00E+00	2	1.30E+09	13.2	immunoglobulin light chain lambda 2	11992186
M462	0	0.00E+00	1	2.86E+08	0	0.00E+00	11.9	hypothetical protein XP_094914	7573285
M463	0	0.00E+00	0	0.00E+00	1	7.22E+07	13.4	anti-DNA immunoglobulin IgG2a heavy chain	2952232
M464	0	0.00E+00	3	3.30E+09	0	0.00E+00	11.3	IG LAMBDA CHAIN V-IV REGION MOL	126570
M465	1	1.03E+08	0	0.00E+00	0	0.00E+00	96.0	thrombospondin 4	14726546
M466	1	1.87E	0	0.00	0	0.00E+	13.3	Ig kappa chain V region	54290

		+08		E+00		00			6
M467	0	0.00E +00	1	3.70 E+08	0	0.00E+ 00	9.4	hypothetical protein XP_099730	18549 142
M468	0	0.00E +00	1	1.67 E+07	0	0.00E+ 00	23.0	Rho GDP dissociation inhibitor (GDI) beta Ly-GDI	10835 002
M469	0	0.00E +00	0	0.00 E+00	1	1.34E+ 08	9.4	immunoglobulin kappa chain variable region	96632 39
M470	0	0.00E +00	1	4.45 E+08	0	0.00E+ 00	8.9	immunoglobulin lambda chain variable region	16075 970
M471	0	0.00E +00	0	0.00 E+00	1	2.04E+ 08	20.5	hypothetical protein XP_106266	18558 469
M472	0	0.00E +00	0	0.00 E+00	3	2.85E+ 08	9.9	immunoglobulin kappa chain variable region	19773 379
M473	2	2.96E +09	0	0.00 E+00	0	0.00E+ 00	12.2	immunoglobulin light chain variable region	19744 488
M474	0	0.00E +00	1	5.07 E+08	0	0.00E+ 00	10.6	kappa-immunoglobulin	30697 4
M475	0	0.00E +00	0	0.00 E+00	1	1.19E+ 09	96.1	Sarcoplasmic reticulum histidine-rich calcium-binding protein precursor	13487 4
M476	1	1.26E +09	0	0.00 E+00	0	0.00E+ 00	10.7	IG KAPPA CHAIN V-I REGION KUE	58338 69
M477	1	0.00E +00	2	7.52 E+08	1	2.28E+ 08	11.5	immunoglobulin lambda chain variable region	99683 86
M478	1	5.80E +08	0	0.00 E+00	0	0.00E+ 00	29.6	similar to unnamed protein product	18544 087
M479	1	6.33E +06	0	0.00 E+00	2	9.94E+ 07	10.0	Ig kappa light chain V-kappa 3 (VJ)	48091 9
M480	0	0.00E +00	2	4.32 E+08	0	0.00E+ 00	12.9	immunoglobulin light chain lambda 3	11992 196
M481	1	1.31E +08	0	0.00 E+00	2	6.65E+ 08	10.8	immunoglobulin kappa chain	72255 2
M482	1	2.80E +08	0	0.00 E+00	0	0.00E+ 00	60.9	similar to MYOSIN VB (MYOSIN 5B)	17481 871
M483	0	0.00E +00	1	2.74 E+08	0	0.00E+ 00	16.5	osteoglycin preproprotein mimecan osteoinductive factor	11279 056
M484	0	0.00E +00	0	0.00 E+00	1	7.53E+ 08	11.7	immunoglobulin lambda chain V-J region	12357 79
M485	1	2.03E +08	0	0.00 E+00	1	6.38E+ 07	11.6	immunoglobulin lambda light chain variable region	31425 10
M486	0	0.00E +00	0	0.00 E+00	1	1.90E+ 08	11.6	immunoglobulin kappa light chain	15616 12
M487	1	1.60E +08	0	0.00 E+00	0	0.00E+ 00	11.6	immunoglobulin kappa light chain variable region	92464 81
M488	0	0.00E +00	2	4.74 E+08	0	0.00E+ 00	3.7	NO_WORTHWHILE_NAMES_FOUND	11178 6
M489	1	2.02E +07	0	0.00 E+00	0	0.00E+ 00	47.9	procollagen C-endopeptidase enhancer procollagen, type 1, COOH-terminal proteinase enhancer	45056 43
M490	0	0.00E +00	0	0.00 E+00	1	3.13E+ 08	10.9	immunoglobulin kappa chain	72261 6
M491	1	4.39E +08	0	0.00 E+00	0	0.00E+ 00	10.4	immunoglobulin kappa chain variable region	72487 31
M492	0	0.00E +00	1	4.82 E+08	0	0.00E+ 00	85.2	G protein-coupled receptor kinase-associated ADP ribosylation factor GTPase-	13929 158

								activating protein (GIT1)	
M493	1	2.81E+08	0	0.00E+00	0	0.00E+00	2.6	NO_WORTHWHILE_NAMES_FOUND	106610
M494	1	1.88E+07	0	0.00E+00	0	0.00E+00	44.7	preprocollagen (AA -22 to 450) (1500 is 1st base in codon)	30016
M495	1	5.72E+08	1	7.45E+08	0	0.00E+00	17.7	hypothetical protein XP_066508	17459719
M496	0	0.00E+00	0	0.00E+00	1	4.79E+07	11.6	CYTOCHROME C	118014
M497	0	0.00E+00	1	2.19E+07	0	0.00E+00	52.3	SELENIUM-BINDING PROTEIN 1	6094240
M498	0	0.00E+00	1	1.16E+08	0	0.00E+00	12.0	IG KAPPA CHAIN V-V REGION HP 91A3	125850
M499	2	7.23E+08	1	4.68E+08	0	0.00E+00	11.4	immunoglobulin lambda chain V-J region	1235777
M500	1	1.19E+07	0	0.00E+00	0	0.00E+00	12.3	Serum amyloid A protein (SAA) [Contains: Amyloid protein A (Amyloid fibril protein AA)]	7531274
M501	0	0.00E+00	1	3.30E+08	0	0.00E+00	15.1	profilin 1	4826898
M502	2	1.81E+08	0	0.00E+00	0	0.00E+00	44.0	hypothetical protein XP_090703	18582665
M503	1	5.48E+08	0	0.00E+00	0	0.00E+00	29.8	hypothetical protein XP_108953	18572548
M504	1	2.62E+08	0	0.00E+00	0	0.00E+00	10.8	immunoglobulin kappa chain	33686
M505	0	0.00E+00	2	3.23E+08	0	0.00E+00	39.3	hypothetical protein XP_068042	18557730
M506	1	1.01E+08	0	0.00E+00	0	0.00E+00	8.8	Ig kappa chain VKIII-JK4	470434
M507	0	0.00E+00	1	1.99E+08	0	0.00E+00	10.2	immunoglobulin lambda light chain	4566033
M508	1	8.61E+08	0	0.00E+00	1	1.70E+08	11.2	immunoglobulin kappa light chain variable region	4378362
M509	0	0.00E+00	0	0.00E+00	1	1.14E+08	23.7	peptidylprolyl isomerase B (cyclophilin B)	4758950
M510	1	9.56E+08	0	0.00E+00	0	0.00E+00	11.2	immunoglobulin lambda-2 variable region	13016680
M511	1	5.60E+06	0	0.00E+00	0	0.00E+00	31.5	myristoylated alanine-rich protein kinase C substrate 80K-L phosphomyristin myristoylated alanine-rich protein kinase C substrate (MARCKS, 80K-L)	11125772
M512	0	0.00E+00	1	7.44E+07	0	0.00E+00	1.8	vascular cell adhesion molecule-1	531882
M513	0	0.00E+00	1	7.18E+07	0	0.00E+00	11.3	IG LAMBDA CHAIN V-IV REGION X	126567
M514	0	0.00E+00	0	0.00E+00	1	2.19E+07	12.1	immunoglobulin light chain variable region	5327155
M515	0	0.00E+00	0	0.00E+00	1	4.61E+08	12.9	lambda 1 immunoglobin light chain variable region	5524097
M516	0	0.00E+00	1	3.64E+07	0	0.00E+00	7.3	FXYD domain-containing ion transport regulator 2, isoform 1 ATPase, Na+/K+ transporting, gamma 1 polypeptide Sodium-	11125766

								potassium-ATPase, gamma polypeptide	
M517	0	0.00E+00	1	1.28E+08	0	0.00E+00	12.7	hypothetical protein XP_092314	2597940
M518	1	5.11E+08	0	0.00E+00	0	0.00E+00	12.6	immunoglobulin light chain variable region	18025698
M519	0	0.00E+00	1	7.72E+07	0	0.00E+00	37.5	hypothetical protein DKFZp434E042 KIAA0732 protein	8922132
M520	0	0.00E+00	0	0.00E+00	1	9.09E+07	12.1	hypothetical protein XP_106837	18561365
M521	1	8.12E+08	0	0.00E+00	0	0.00E+00	11.5	immunoglobulin light chain variable region	6735442
M522	0	0.00E+00	1	1.51E+07	0	0.00E+00	64.2	NO_WORTHWHILE_NAMES_FOUND	16041088
M523	0	0.00E+00	1	1.51E+07	0	0.00E+00	14.1	hypothetical protein XP_071142	17452896
M524	0	0.00E+00	1	1.52E+07	0	0.00E+00	14.0	Ig heavy chain variable region	951291
M525	0	0.00E+00	1	3.24E+09	1	2.64E+07	25.2	hypothetical protein XP_095665	18571355
M526	0	0.00E+00	0	0.00E+00	1	2.80E+08	91.1	hypothetical protein DKFZp761I241	14165290
M527	0	0.00E+00	0	0.00E+00	1	5.69E+07	11.9	protein LOC,Bence-Jones	223970
M528	1	7.47E+08	0	0.00E+00	0	0.00E+00	43.1	farnesyl-protein transferase beta-subunit	2135098
M529	1	1.17E+08	0	0.00E+00	0	0.00E+00	16.2	hypothetical protein XP_103646	18550923
M530	1	6.32E+08	0	0.00E+00	0	0.00E+00	10.1	immunoglobulin kappa chain	11137029
M531	0	0.00E+00	1	7.99E+08	0	0.00E+00	42.0	Burkitt lymphoma receptor 1, isoform 1 Burkitt lymphoma receptor 1, GTP-binding protein C-X-C chemokine receptor type 5 monocyte-derived receptor 15	4502415
M532	0	0.00E+00	1	2.63E+08	0	0.00E+00	94.6	KIAA1817 protein	14763137
M533	1	3.20E+07	0	0.00E+00	0	0.00E+00	50.2	fibulin 5	14748759
M534	0	0.00E+00	1	2.36E+09	0	0.00E+00	11.9	immunoglobulin light chain variable region	18698379
M535	0	0.00E+00	0	0.00E+00	1	1.69E+08	11.5	immunoglobulin lambda light chain variable region	12830385
M536	1	7.05E+07	0	0.00E+00	0	0.00E+00	44.0	Ig superfamily protein	6005958
M537	0	0.00E+00	1	5.90E+08	0	0.00E+00	11.9	Chain A, Structural Comparison Of Amyloidogenic Light Chain Dimer In Two Crystal Forms With Nonamyloidogenic Counterparts	2135446
M538	1	1.53E+08	0	0.00E+00	0	0.00E+00	17.2	IgM	1399519
M539	0	0.00E+00	1	2.86E+08	0	0.00E+00	77.0	VILLIN-LIKE PROTEIN	14548297

M540	0	0.00E+00	1	2.62E+06	0	0.00E+00	25.3	hypothetical protein XP_086067	18591092
M541	0	0.00E+00	0	0.00E+00	1	1.74E+10	6.0	beta-A3 crystallin	2338452
M542	0	0.00E+00	0	0.00E+00	1	5.01E+08	9.4	immunoglobulin lambda chain variable region	16076054
M543	0	0.00E+00	1	9.24E+07	0	0.00E+00	72.6	chromosome 1 open reading frame 14 GE36 gene	12017952
M544	1	1.29E+08	0	0.00E+00	0	0.00E+00	18.2	immunoglobulin kappa-chain	12750740
M545	0	0.00E+00	1	4.23E+07	0	0.00E+00	43.1	nuclear receptor coactivator 4 RFG	12852362
M546	0	0.00E+00	1	2.16E+08	0	0.00E+00	32.4	hypothetical protein XP_090449	18580746
M547	0	0.00E+00	1	4.37E+07	0	0.00E+00	24.4	hypothetical protein XP_067233	17435366
M548	1	1.93E+08	0	0.00E+00	0	0.00E+00	11.4	anti-HCV E2 antibody, VK segment	4837695
M549	1	1.11E+07	0	0.00E+00	0	0.00E+00	13.7	immunoglobulin mu heavy chain variable region	4995318
M550	1	3.31E+08	0	0.00E+00	0	0.00E+00	85.6	squamous cell carcinoma antigen recognized by T-cells	2342526
M551	0	0.00E+00	0	0.00E+00	1	6.72E+07	9.7	immunoglobulin lambda light chain variable region	18041866
M552	513	5.69E+11	344	3.64E+11	618	7.54E+11	192.8	progesterone-induced blocking factor 1	4502501
M553	11	1.20E+09	10	7.03E+08	35	4.55E+09	515.3	apolipoprotein B fragment	28780
M554	36	1.44E+10	26	9.03E+09	51	2.63E+10	139.1	H factor (complement)-like 3 factor H-related gene 2	4504375
M555	81	4.85E+10	55	2.19E+10	103	5.75E+10	103.4	PRO1851	13432192
M556	53	4.97E+10	40	3.33E+10	60	3.28E+10	72.0	KININOGEN PRECURSOR (ALPHA-2-THIOL PROTEINASE INHIBITOR) [CONTAINS: BRADYKININ]	386852
M557	133	4.19E+11	109	2.98E+11	131	4.99E+11	34.3	alpha-2-glycoprotein 1, zinc Alpha-2-glycoprotein, zinc	4502337
M558	18	1.02E+10	19	8.53E+09	20	1.30E+10	80.2	Complement C1r component precursor	115204
M559	17	8.83E+08	0	0.00E+00	0	0.00E+00	881.8	NO_WORTHWHILE_NAMES_FOUND	27717403
M560	60	7.58E+10	48	5.21E+10	64	6.87E+10	51.9	NO_WORTHWHILE_NAMES_FOUND	69990
M561	16	2.73E+09	9	1.09E+09	26	7.35E+09	101.4	inter-alpha-trypsin inhibitor C-terminal	4504781
M562	15	4.51E+09	9	4.38E+09	25	1.06E+10	106.4	NO_WORTHWHILE_NAMES_FOUND	125000
M563	37	4.47E+10	25	1.38E+10	62	8.18E+10	65.0	Unknown (protein for MGC:23134)	16041759
M564	72	1.24E+11	78	1.20E+11	42	8.61E+10	38.2	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 13 (RNA helicase A)	21707947
M565	17	4.97E+09	20	4.21E+09	26	1.30E+10	68.0	peptidoglycan recognition protein-like TAG-like	21361845

M566	203	3.08E+11	329	4.25E+11	1742	2.13E+12	15.3	alpha 2 globin	13650074
M567	8	1.83E+09	11	5.69E+08	2	1.04E+08	58.8	Keratin, type I cytoskeletal 10 (Cytokeratin 10) (K10) (CK 10)	21961605
M568	29	1.52E+10	35	2.63E+10	19	1.47E+10	46.3	serine (or cysteine) proteinase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 1 pigment epithelium-derived factor	20178323
M569	80	1.46E+11	88	9.06E+10	22	3.28E+09	13.5	serum amyloid A1	36308
M570	8	5.03E+09	13	1.23E+10	4	3.26E+09	43.8	sex hormone-binding globulin Sex hormone-binding globulin (androgen binding protein)	338075
M571	52	4.78E+10	46	5.88E+10	119	1.74E+11	32.9	NO_WORTHWHILE_NAMES_FOUND	27754776
M572	161	1.26E+11	125	6.02E+10	332	3.48E+11	26.5	Complement C1q subcomponent, B chain precursor	399140
M573	35	7.75E+10	26	1.83E+10	39	3.44E+10	25.7	NO_WORTHWHILE_NAMES_FOUND	27363488
M574	4	5.78E+08	0	0.00E+00	9	1.73E+09	21.9	peroxiredoxin 2 thioredoxin-dependent peroxide reductase 1 (thiol-specific antioxidant 1, natural killer-enhancing factor B) thiol-specific antioxidant 1 natural killer-enhancing factor B	2507169
M575	0	0.00E+00	6	6.35E+08	0	0.00E+00	22.5	NO_WORTHWHILE_NAMES_FOUND	27672700
M576	2	2.95E+08	0	0.00E+00	8	2.55E+09	20.6	MBL-associated serine protease(MASP)-2	21264361
M577	6	1.93E+09	3	1.21E+08	0	0.00E+00	10.9	NO_WORTHWHILE_NAMES_FOUND	29888
M578	0	0.00E+00	5	9.87E+08	5	8.42E+08	553.1	NO_WORTHWHILE_NAMES_FOUND	13876386
M579	4	1.48E+09	1	1.25E+08	1	2.82E+07	41.6	actin, gamma 2, smooth muscle, enteric	71621
M580	7	4.76E+09	5	8.79E+08	13	9.45E+09	10.5	INSULIN-LIKE GROWTH FACTOR II PRECURSOR (IGF-II) (ERYTHROTROPIN)	2136466
M581	2	5.54E+08	0	0.00E+00	4	3.36E+08	104.8	NO_WORTHWHILE_NAMES_FOUND	4559406
M582	1	3.91E+08	0	0.00E+00	4	1.61E+09	48.5	NO_WORTHWHILE_NAMES_FOUND	21361302
M583	2	5.90E+08	0	0.00E+00	1	6.21E+07	154.0	endostatin variant	7717447
M584	5	1.92E+09	4	6.88E+08	6	1.29E+09	21.3	HSPC336	22091452
M585	2	2.53E+09	3	1.27E+09	7	4.17E+09	11.6	monoclonal antibody MAX5 immunoglobulin light chain variable region	2995683
M586	4	1.15E+09	0	0.00E+00	5	8.10E+08	138.1	NO_WORTHWHILE_NAMES_FOUND	1362867
M587	0	0.00E	0	0.00	3	2.97E+	414.9	NO_WORTHWHILE_NA	27713

		+00		E+00		07		MES_FOUND	050
M588	2	1.96E +08	0	0.00 E+00	3	3.10E+ 08	22.3	Superoxide dismutase [Mn], mitochondrial precursor	23503 532
M589	0	0.00E +00	2	3.13 E+08	2	9.96E+ 08	25.9	superoxide dismutase 3, extracellular	45071 51
M590	3	7.18E +09	1	8.13 E+08	2	6.38E+ 07	129.4	thrombospondin-1	45074 85
M591	2	2.32E +08	0	0.00 E+00	0	0.00E+ 00	111.0	ataxin-2 related protein	38204 84
M592	2	6.09E +08	0	0.00 E+00	0	0.00E+ 00	37.1	calumenin precursor	14718 453
M593	3	1.20E +09	0	0.00 E+00	1	7.40E+ 07	29.2	carbonic anhydrase II	15080 386
M594	0	0.00E +00	2	1.47 E+07	0	0.00E+ 00	13.3	dolichyl-phosphate mannosyltransferase polypeptide 3	19424 120
M595	0	0.00E +00	2	6.64 E+08	0	0.00E+ 00	26.2	Fc fragment of IgG, low affinity IIIb, receptor for (CD16)	10835 139
M596	0	0.00E +00	0	0.00 E+00	3	1.69E+ 09	80.4	HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN PRECURSOR (MACROPHAGE STIMULATORY PROTEIN) (MSP) (MACROPHAGE STIMULATING PROTEIN)	12311 4
M597	3	2.14E +08	0	0.00 E+00	0	0.00E+ 00	15.2	INSULIN-LIKE GROWTH FACTOR IB PRECURSOR (IGF-IB) (SOMATOMEDIN C)	10673 9
M598	0	0.00E +00	2	6.73 E+07	0	0.00E+ 00	27.7	keratin	77172 38
M599	0	0.00E +00	0	0.00 E+00	2	1.04E+ 09	254.4	KIAA0539 protein	27501 097
M600	2	1.01E +09	0	0.00 E+00	0	0.00E+ 00	252.8	KIAA1241 protein	21281 669
M601	0	0.00E +00	0	0.00 E+00	2	1.09E+ 08	83.9	myeloperoxidase	45577 59
M602	0	0.00E +00	0	0.00 E+00	3	5.15E+ 08	274.6	NO_WORTHWHILE_NAMES_FOUND	45071 95
M603	2	2.46E +09	0	0.00 E+00	0	0.00E+ 00	101.4	NO_WORTHWHILE_NAMES_FOUND	67543 66
M604	0	0.00E +00	0	0.00 E+00	3	2.06E+ 09	201.4	NO_WORTHWHILE_NAMES_FOUND	67554 06
M605	0	0.00E +00	0	0.00 E+00	3	1.47E+ 08	469.4	NO_WORTHWHILE_NAMES_FOUND	74275 17
M606	0	0.00E +00	2	1.85 E+08	0	0.00E+ 00	313.5	NO_WORTHWHILE_NAMES_FOUND	75497 81
M607	0	0.00E +00	0	0.00 E+00	2	8.35E+ 08	250.8	NO_WORTHWHILE_NAMES_FOUND	76619 60
M608	0	0.00E +00	0	0.00 E+00	2	3.76E+ 08	104.4	NO_WORTHWHILE_NAMES_FOUND	13376 091
M609	0	0.00E +00	3	3.77 E+08	0	0.00E+ 00	541.3	NO_WORTHWHILE_NAMES_FOUND	14335 446
M610	0	0.00E +00	0	0.00 E+00	5	8.13E+ 09	99.2	NO_WORTHWHILE_NAMES_FOUND	19923 540

M611	1	1.91E+08	1	2.38E+08	4	1.47E+09	81.9	NO_WORTHWHILE_NAMES_FOUND	21264359
M612	16	7.59E+09	0	0.00E+00	1	2.55E+09	26.9	NO_WORTHWHILE_NAMES_FOUND	21450229
M613	0	0.00E+00	1	3.60E+08	4	2.52E+09	11.5	NO_WORTHWHILE_NAMES_FOUND	21669273
M614	0	0.00E+00	0	0.00E+00	2	1.10E+08	75.5	NO_WORTHWHILE_NAMES_FOUND	21752646
M615	0	0.00E+00	0	0.00E+00	2	4.70E+08	198.7	NO_WORTHWHILE_NAMES_FOUND	22094095
M616	0	0.00E+00	0	0.00E+00	2	7.81E+08	909.3	NO_WORTHWHILE_NAMES_FOUND	27699130
M617	0	0.00E+00	2	6.83E+08	0	0.00E+00	512.3	NO_WORTHWHILE_NAMES_FOUND	27729601
M618	1	2.93E+08	3	9.53E+07	1	2.75E+07	33.9	osteoglycin preproprotein mimecan osteoinductive factor	7661704
M619	0	0.00E+00	2	3.11E+08	0	0.00E+00	41.9	pregnancy-associated glycoprotein-2	28603724
M620	0	0.00E+00	0	0.00E+00	4	8.61E+08	88.0	suppressor of var1, 3-like 1 (S. cerevisiae) suppressor of var1 (S.cerevisiae) 3-like 1	4507315
M621	0	0.00E+00	2	2.48E+07	0	0.00E+00	50.8	zyxin	16877914
M622	0	0.00E+00	1	2.48E+07	0	0.00E+00	48.4		4503005
M623	0	0.00E+00	0	0.00E+00	1	1.06E+07	25.5		15451786
M624	0	0.00E+00	0	0.00E+00	1	5.14E+08	11.9	amyloid lambda 6 light chain variable region NEG	14279403
M625	0	0.00E+00	0	0.00E+00	1	1.30E+08	5.8	beta-chemokine RANTES precursor	25991891
M626	1	7.34E+08	0	0.00E+00	0	0.00E+00	35.7	bone marrow stromal cell antigen 1 precursor	15082365
M627	1	1.10E+07	0	0.00E+00	0	0.00E+00	61.7	bridging integrator 2 bridging integrator-2 breast cancer associated protein BRAP1	7705296
M628	1	9.73E+07	0	0.00E+00	0	0.00E+00	9.7	calmodulin 2 (phosphorylase kinase, delta) phosphorylase kinase delta	2654179
M629	0	0.00E+00	1	8.45E+07	1	8.11E+07	15.9	calmodulin-like skin protein	24657605
M630	0	0.00E+00	1	8.66E+06	0	0.00E+00	25.8	carbonyl reductase	12804319
M631	0	0.00E+00	1	2.69E+08	0	0.00E+00	76.6	CD44 ANTIGEN PRECURSOR (PHAGOCYTIC GLYCOPROTEIN I) (PGP-1) (HUTCH-I) (EXTRACELLULAR MATRIX RECEPTOR-III) (ECMR-III) (GP90 LYMPHOCYTE HOMING/ADHESION RECEPTOR) (HERMES ANTIGEN) (HYALURONATE	2135073

								RECEPTOR)	
M632	0	0.00E+00	0	0.00E+00	1	3.61E+07	48.0	complement C1r-like proteinase precursor,	10436374
M633	0	0.00E+00	0	0.00E+00	1	1.70E+07	51.0	coronin, actin binding protein, 1A coronin, actin-binding, 1A coronin, actin-binding protein, 1A	5902134
M634	1	6.24E+07	0	0.00E+00	0	0.00E+00	7.3	defensin, beta 1, preproprotein beta defensin 1	14486454
M635	0	0.00E+00	0	0.00E+00	1	6.36E+08	51.6	FK506 binding protein 4 (59 kDa)	6753882
M636	1	3.84E+08	1	1.31E+08	0	0.00E+00	77.5	FLJ00033 protein	22748647
M637	1	2.13E+08	0	0.00E+00	0	0.00E+00	65.3	galectin 3 binding protein L3 antigen Mac-2-binding protein serum protein 90K	5031863
M638	1	3.31E+08	2	4.22E+08	1	2.49E+08	22.7	gene trap locus 3	26353782
M639	0	0.00E+00	1	4.03E+07	0	0.00E+00	16.0	Hemoglobin beta fetal chain (Hemoglobin gamma chain)	27819608
M640	0	0.00E+00	0	0.00E+00	1	7.59E+08	8.8	hypothetical protein XP_066346	21310689
M641	0	0.00E+00	0	0.00E+00	1	7.34E+06	57.8	intercellular adhesion molecule 1 precursor CD54	4557878
M642	0	0.00E+00	1	7.38E+07	1	2.63E+08	30.7	intercellular adhesion molecule 2 precursor	386792
M643	1	1.18E+08	0	0.00E+00	0	0.00E+00	65.4	interleukin 1 receptor accessory protein	27902526
M644	0	0.00E+00	0	0.00E+00	1	1.81E+09	39.7	isocitrate dehydrogenase 3 (NAD+) alpha	26339056
M645	0	0.00E+00	1	5.21E+06	0	0.00E+00	51.7	keratin 14 cytokeratin 14	12803709
M646	1	3.27E+08	0	0.00E+00	0	0.00E+00	93.1	KIAA1061 protein	22049346
M647	0	0.00E+00	1	8.33E+07	1	8.97E+07	84.4	meprin A, alpha (PABA peptide hydrolase) Meprin A, alpha	12141249
M648	0	0.00E+00	0	0.00E+00	1	7.30E+08	4.6	microfibril-associated glycoprotein MAP	545599
M649	0	0.00E+00	0	0.00E+00	4	4.27E+09	11.6	NO_WORTHWHILE_NAMES_FOUND	542909
M650	0	0.00E+00	0	0.00E+00	1	1.01E+08	60.9	NO_WORTHWHILE_NAMES_FOUND	3327120
M651	1	4.19E+07	0	0.00E+00	0	0.00E+00	105.8	NO_WORTHWHILE_NAMES_FOUND	4507489
M652	0	0.00E+00	3	1.10E+09	0	0.00E+00	138.1	NO_WORTHWHILE_NAMES_FOUND	19352987
M653	0	0.00E+00	0	0.00E+00	1	1.54E+07	36.8	NO_WORTHWHILE_NAMES_FOUND	20149646
M654	0	0.00E+00	0	0.00E+00	1	7.53E+08	12.1	NO_WORTHWHILE_NAMES_FOUND	20372508
M655	0	0.00E+00	0	0.00E+00	1	1.47E+08	14.4	NO_WORTHWHILE_NAMES_FOUND	20850402
M656	0	0.00E	0	0.00	1	1.26E+	53.7	NO_WORTHWHILE_NA	20871

		+00		E+00		08		MES_FOUND	931
M657	3	1.76E +10	1	1.67 E+09	0	0.00E+ 00	12.9	NO_WORTHWHILE_NA MES_FOUND	20892 231
M658	1	1.61E +09	0	0.00 E+00	0	0.00E+ 00	50.0	NO_WORTHWHILE_NA MES_FOUND	20898 918
M659	0	0.00E +00	0	0.00 E+00	2	1.17E+ 09	26.0	NO_WORTHWHILE_NA MES_FOUND	21410 817
M660	1	1.59E +09	0	0.00 E+00	1	1.36E+ 09	48.3	NO_WORTHWHILE_NA MES_FOUND	21687 104
M661	0	0.00E +00	1	1.01 E+08	0	0.00E+ 00	107.1	NO_WORTHWHILE_NA MES_FOUND	23273 447
M662	1	7.50E +08	0	0.00 E+00	0	0.00E+ 00	70.8	NO_WORTHWHILE_NA MES_FOUND	26335 149
M663	1	6.68E +08	0	0.00 E+00	0	0.00E+ 00	41.4	NO_WORTHWHILE_NA MES_FOUND	26342 200
M664	1	9.69E +08	0	0.00 E+00	0	0.00E+ 00	117.9	NO_WORTHWHILE_NA MES_FOUND	27503 142
M665	1	2.07E +09	0	0.00 E+00	0	0.00E+ 00	52.5	NO_WORTHWHILE_NA MES_FOUND	27660 044
M666	0	0.00E +00	1	4.03 E+07	0	0.00E+ 00	15.2	NO_WORTHWHILE_NA MES_FOUND	27668 422
M667	1	4.38E +08	0	0.00 E+00	0	0.00E+ 00	47.4	NO_WORTHWHILE_NA MES_FOUND	27670 153
M668	0	0.00E +00	1	4.52 E+07	0	0.00E+ 00	84.7	NO_WORTHWHILE_NA MES_FOUND	27680 380
M669	0	0.00E +00	1	1.66 E+08	0	0.00E+ 00	264.0	NO_WORTHWHILE_NA MES_FOUND	27691 330
M670	0	0.00E +00	1	1.25 E+08	0	0.00E+ 00	12.3	NO_WORTHWHILE_NA MES_FOUND	27708 240
M671	1	2.04E +08	0	0.00 E+00	1	8.72E+ 07	20.3	NO_WORTHWHILE_NA MES_FOUND	27894 364
M672	0	0.00E +00	0	0.00 E+00	1	4.64E+ 09	56.2	NO_WORTHWHILE_NA MES_FOUND	28511 899
M673	0	0.00E +00	1	4.29 E+06	0	0.00E+ 00	10.3	proline-rich acidic protein	27358 010
M674	0	0.00E +00	0	0.00 E+00	1	1.73E+ 07	40.4	RIKEN cDNA 1300009F09	26350 553
M675	0	0.00E +00	0	0.00 E+00	1	3.14E+ 07	17.9	sonic hedgehog homolog (Drosophila) Sonic hedgehog (Drosophila), human homolog of sonic hedgehog (Drosophila) homolog	55663 17
M676	0	0.00E +00	0	0.00 E+00	1	4.87E+ 07	22.4	transgelin 2 SM22-alpha homolog	20830 019
M677	0	0.00E +00	1	2.03 E+08	0	0.00E+ 00	44.9	unknown	22748 639
M678	0	0.00E +00	1	9.72 E+07	0	0.00E+ 00	62.3	Unknown (protein for IMAGE:3953315)	14043 412
M679	0	0.00E +00	1	1.44 E+08	0	0.00E+ 00	78.3	unnamed protein product	96510 75

TABLE 2

		Erosive Serum		Non Erosive Serum		Normal Serum							
Marker #	**Gene name**	**# spectra**	**Total intensity**	**# spectra**	**Total intensity**	**# spectra**	**Total intensity**	**E:N**	**E:H**	**mw (kDa)**	**acces s. no.**	**SEQ ID NO (nts)**	**SEQ ID NO (aa)**
M108	lysozyme C (1,4-beta-N-acetylmuramidase C)	19	9.56E+09	14	7.61E+09			**1:1**	**0 in H**	16.5	4557894	1	2
M105	lumican (keratan sulfate proteoglycan)	16				2	7.69E+08	**0 in N**	**25:1**	38.4	1708878	3	4
M90	angiogenin (ribonuclease 5)	21	3.45E+09	15	2.37E+09			**1:1**	**0 in H**	16.6	18307851	5	6
M129	ribonuclease 4	7	1.49E+09	10	1.48E+09			**1:1**	**0 in H**	13.8	4506557	7	8
M99	Platelet factor 4 variant precursor (PF4VAR1)	54		18	3.83E+09	10	1.69E+09	**20:1**	**2:1**	11.6	4505735	9	10
M168	preproteolysin	2	1.29E+08	3	8.73E+07			**2:1**	**0 in H**	11.3	16751921	11	12
M178	ficolin 2 (ficolin B, serum lectin P35)	11		1	8.65E+08	3	9.81E+08	**17:1**	**15:1**	34	4758348	13	14
M191	Insulin-like growth factor binding protein 5	2	1.98E+08	3	3.66E+08			**1:2**	**0 in H**	30.6	10834982	15	16
M256	secretory granule proteoglycan core protein	2	1.20E+08	1	2.83E+08			**1:2**	**0 in H**	17.6	4506045	17	18

M145	small inducible cytokine A14 (HCC-1/HCC-3)	8	2.38E+09	2	7.60E+07			30:1	0 in H	10.7	47590 70	19	20
M239	small inducible cytokine A16 (HCC-4)	2	4.68E+07	1	9.71E+07			1:2	0 in H	13.6	47590 74	21	22
M185	small inducible cytokine A18 (MIP-4)	6	5.48E+09	4	4.18E+08			13:1	0 in H	9.8	45068 31	23	24
M266	fractalkine (small inducible cytokine D1)	2	1.77E+08					0 in N	0 in H	42.2	45068 57	25	26
M300	stromal cell-derived factor 1 (CXCL12)	1	1.50E+08	1	3.50E+07			4:1	0 in H	10	10834 988	27	28
M186	collagen alpha 1(XVIII) chain	3	1.08E+09					0 in N	0 in H	70.4	29205 35	29	30
M220	dJ797M17.1 (Dermatopontin)	2	4.33E+08	1	7.52E+07			5:1	0 in H	24	14736 977	31	32
M73	serum amyloid A2	208		38	9.92E+09	24	5.80E+09	39:1	67:1	11.6	13540 475	33	34
M100	S100 calcium-binding protein A8 cystic fibrosis antigen calgranulin A	21	3.53E+09	9	4.93E+08	4	2.01E+08	7:1	17:1	10.9	45067 71	35	36
M97	S100 calcium-binding	29	7.29E+09	15	2.69E+09	9	9.80E+08	3:1	7:1	13.2	45067 73	37	38

	protein A9 calgranulin B												
M264	S100 A12 protein, calgranulin C	2	4.06E+08	1	7.17E+07			5:1	0 in H	10.4	21469 72	39	40
M196	Osteopontin precursor (bone sialoprotein 1) (secreted phosphoprotei n 1) (SPP-1) (Nephroponti n) (uropontin)	5	5.50E+08					0 in N	0 in H	35.4	21197 10	41	42
M590	thrombospond in-1	3	7.18E+09	1	8.13E+08	2	6.38E+07	9:1	112:1	129. 4	45074 85	43	44
M592	calumenin precursor	2	3.09E+08					0 in N	0 in H	37.1	14718 453	45	46
M593	carbonic anhydrase II	3	1.20E+09			1	7.40E+07	0 in N	16:1	29.2	15080 386	47	48
M618	osteoglycin preproprotein mimecan osteoinductiv e factor	1	2.93E+08	3	9.53E+07	1	2.75E+07	3:1	10:1	33.9	76617 04	49	50
M657	NO_WORTH WHILE_NA MES_FOUN D	3		1	1.67E+09			10:1	0 in H	12.9	20892 231	51	52

Other Embodiments

[00307] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

[00308] The contents of all references, patents, published patent applications, and database records cited throughout this application are hereby incorporated by reference.

What is claimed is:

1. A method of assessing whether a patient is afflicted with RA, the method comprising:

a) determining the level of expression of one or more markers in a patient sample, wherein the one or more markers are selected from the group consisting of markers listed in Table 1;

b) determining the normal level of expression of the one or more markers in a control sample; and

c) comparing the level of expression of the one or more markers in the patient sample to the level of expression of the one or more markers in the control sample,

wherein a significant difference in the level of expression of the one or more markers in the patient sample compared to the normal level is an indication that the patient is afflicted with RA.

2. The method of claim 1, wherein the level of expression is determined by detecting the amount of marker protein present in the sample.

3. The method of claim 1, wherein the level of expression is determined by detecting the amount of mRNA that encodes a marker protein present in the sample.

4. A method of assessing whether a patient is afflicted with RA, the method comprising:

a) determining the level of expression of a plurality of markers in a patient sample, wherein at least one of the markers is selected from Table 2;

b) determining the normal level of expression of the plurality of markers in a control sample; and

c) comparing the level of expression of the plurality of markers in the patient sample to the level of expression of the plurality of markers in the control sample,

wherein a significant difference in the level of expression of the plurality of markers in the patient sample compared to the normal level is an indication that the patient is afflicted with RA.

5. The method of claim 4, wherein the control is the level of expression of the one or more markers in a non-erosive RA patient sample.

6. The method of claim 4, wherein the level of expression is determined by detecting the amount of marker protein present in the sample.

7. The method of claim 4, wherein the level of expression is determined by detecting the amount of mRNA that encodes a marker protein present in the sample.

8. A method of assessing whether a patient is afflicted with erosive RA, the method comprising:

a) determining the level of expression of one or more markers in a patient sample, wherein the one or more markers are selected from the group consisting of markers listed in Table 2;

b) determining the level of expression of the one or more markers in a control sample; and

c) comparing the level of expression of the one or more markers in the patient sample to the level of expression of the one or more markers in the control sample,

wherein a significant difference between the level of expression of the one or more markers in the patient sample and the control is an indication that the patient is afflicted with erosive RA.

9. The method of claim 5, wherein the control is the level of expression of the one or more markers in a non-erosive RA patient sample.

10. The method of claim 5, wherein the level of expression is determined by detecting the amount of marker protein present in the sample.

11. The method of claim 5, wherein the level of expression is determined by detecting the amount of mRNA that encodes a marker protein present in the sample.

12. A method of assessing whether a patient is afflicted with RA, the method comprising determining the level of expression of one or more markers in a patient sample, wherein the one or more markers are selected from the group consisting of markers listed

in Table 1; and wherein a significant difference in the level of expression of the one or more markers in the patient sample compared to a reference sample is an indication that the patient is afflicted with RA.

SEQUENCE LISTING

<110> Millennium Pharmaceuticals, Inc.
Guild, Braydon C.
Liao, Hua
Jones, Michael D.
Wu, Jiang
Zolg, Johannes W.

<120> Compositions, Kits, and Methods for Identification, Assessment, Prevention, and Therapy of Rheumatoid Arthritis

<130> MPI03-037

<150> 60/472,330
<151> 2003-05-21

<160> 52

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 447
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(447)

<400> 1

```
atg aag gct ctc att gtt ctg ggg ctt gtc ctc ctt tct gtt acg gtc   48
Met Lys Ala Leu Ile Val Leu Gly Leu Val Leu Leu Ser Val Thr Val
 1               5                  10                  15

cag ggc aag gtc ttt gaa agg tgt gag ttg gcc aga act ctg aaa aga   96
Gln Gly Lys Val Phe Glu Arg Cys Glu Leu Ala Arg Thr Leu Lys Arg
             20                  25                  30

ttg gga atg gat ggc tac agg gga atc agc cta gca aac tgg atg tgt   144
Leu Gly Met Asp Gly Tyr Arg Gly Ile Ser Leu Ala Asn Trp Met Cys
         35                  40                  45

ttg gcc aaa tgg gag agt ggt tac aac aca cga gct aca aac tac aat   192
Leu Ala Lys Trp Glu Ser Gly Tyr Asn Thr Arg Ala Thr Asn Tyr Asn
     50                  55                  60

gct gga gac aga agc act gat tat ggg ata ttt cag atc aat agc cgc   240
Ala Gly Asp Arg Ser Thr Asp Tyr Gly Ile Phe Gln Ile Asn Ser Arg
 65                  70                  75                  80

tac tgg tgt aat gat ggc aaa acc cca gga gca gtt aat gcc tgt cat   288
Tyr Trp Cys Asn Asp Gly Lys Thr Pro Gly Ala Val Asn Ala Cys His
                 85                  90                  95

tta tcc tgc agt gct ttg ctg caa gat aac atc gct gat gct gta gct   336
Leu Ser Cys Ser Ala Leu Leu Gln Asp Asn Ile Ala Asp Ala Val Ala
            100                 105                 110

tgt gca aag agg gtt gtc cgt gat cca caa ggc att aga gca tgg gtg   384
```

```
Cys Ala Lys Arg Val Val Arg Asp Pro Gln Gly Ile Arg Ala Trp Val
        115                 120                 125

gca tgg aga aat cgt tgt caa aac aga gat gtc cgt cag tat gtt caa   432
Ala Trp Arg Asn Arg Cys Gln Asn Arg Asp Val Arg Gln Tyr Val Gln
    130                 135                 140

ggt tgt gga gtg taa                                           447
Gly Cys Gly Val  *
145

<210> 2
<211> 148
<212> PRT
<213> human

<400> 2
Met Lys Ala Leu Ile Val Leu Gly Leu Val Leu Leu Ser Val Thr Val
 1               5                  10                  15
Gln Gly Lys Val Phe Glu Arg Cys Glu Leu Ala Arg Thr Leu Lys Arg
            20                  25                  30
Leu Gly Met Asp Gly Tyr Arg Gly Ile Ser Leu Ala Asn Trp Met Cys
        35                  40                  45
Leu Ala Lys Trp Glu Ser Gly Tyr Asn Thr Arg Ala Thr Asn Tyr Asn
    50                  55                  60
Ala Gly Asp Arg Ser Thr Asp Tyr Gly Ile Phe Gln Ile Asn Ser Arg
65                  70                  75                  80
Tyr Trp Cys Asn Asp Gly Lys Thr Pro Gly Ala Val Asn Ala Cys His
                85                  90                  95
Leu Ser Cys Ser Ala Leu Leu Gln Asp Asn Ile Ala Asp Ala Val Ala
            100                 105                 110
Cys Ala Lys Arg Val Val Arg Asp Pro Gln Gly Ile Arg Ala Trp Val
        115                 120                 125
Ala Trp Arg Asn Arg Cys Gln Asn Arg Asp Val Arg Gln Tyr Val Gln
    130                 135                 140
Gly Cys Gly Val
145

<210> 3
<211> 1017
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(1017)

<400> 3
atg agt cta agt gca ttt act ctc ttc ctg gca ttg att ggt ggt acc   48
Met Ser Leu Ser Ala Phe Thr Leu Phe Leu Ala Leu Ile Gly Gly Thr
 1               5                  10                  15

agt ggc cag tac tat gat tat gat ttt ccc cta tca att tat ggg caa   96
Ser Gly Gln Tyr Tyr Asp Tyr Asp Phe Pro Leu Ser Ile Tyr Gly Gln
            20                  25                  30

tca tca cca aac tgt gca cca gaa tgt aac tgc cct gaa agc tac cca   144
Ser Ser Pro Asn Cys Ala Pro Glu Cys Asn Cys Pro Glu Ser Tyr Pro
        35                  40                  45

agt gcc atg tac tgt gat gag ctg aaa ttg aaa agt gta cca atg gtg   192
```

```
Ser Ala Met Tyr Cys Asp Glu Leu Lys Leu Lys Ser Val Pro Met Val
     50                  55                  60

cct cct gga atc aag tat ctt tac ctt agg aat aac cag att gac cat   240
Pro Pro Gly Ile Lys Tyr Leu Tyr Leu Arg Asn Asn Gln Ile Asp His
 65                  70                  75                  80

att gat gaa aag gcc ttt gag aat gta act gat ctg cag tgg ctc att   288
Ile Asp Glu Lys Ala Phe Glu Asn Val Thr Asp Leu Gln Trp Leu Ile
                85                  90                  95

cta gat cac aac ctt cta gaa aac tcc aag ata aaa ggg aga gtt ttc   336
Leu Asp His Asn Leu Leu Glu Asn Ser Lys Ile Lys Gly Arg Val Phe
            100                 105                 110

tct aaa ttg aaa caa ctg aag aag ctg cat ata aac cac aac aac ctg   384
Ser Lys Leu Lys Gln Leu Lys Lys Leu His Ile Asn His Asn Asn Leu
        115                 120                 125

aca gag tct gtg ggc cca ctt ccc aaa tct ctg gag gat ctg cag ctt   432
Thr Glu Ser Val Gly Pro Leu Pro Lys Ser Leu Glu Asp Leu Gln Leu
    130                 135                 140

act cat aac aag atc aca aag ctg ggc tct ttt gaa gga ttg gta aac   480
Thr His Asn Lys Ile Thr Lys Leu Gly Ser Phe Glu Gly Leu Val Asn
145                 150                 155                 160

ctg acc ttc atc cat ctc cag cac aat cgg ctg aaa gag gat gct gtt   528
Leu Thr Phe Ile His Leu Gln His Asn Arg Leu Lys Glu Asp Ala Val
                165                 170                 175

tca gct gct ttt aaa ggt ctt aaa tca ctc gaa tac ctt gac ttg agc   576
Ser Ala Ala Phe Lys Gly Leu Lys Ser Leu Glu Tyr Leu Asp Leu Ser
            180                 185                 190

ttc aat cag ata gcc aga ctg cct tct ggt ctc cct gtc tct ctt cta   624
Phe Asn Gln Ile Ala Arg Leu Pro Ser Gly Leu Pro Val Ser Leu Leu
        195                 200                 205

act ctc tac tta gac aac aat aag atc agc aac atc cct gat gag tat   672
Thr Leu Tyr Leu Asp Asn Asn Lys Ile Ser Asn Ile Pro Asp Glu Tyr
    210                 215                 220

ttc aag cgt ttt aat gca ttg cag tat ctg cgt tta tct cac aac gaa   720
Phe Lys Arg Phe Asn Ala Leu Gln Tyr Leu Arg Leu Ser His Asn Glu
225                 230                 235                 240

ctg gct gat agt gga ata cct gga aat tct ttc aat gtg tca tcc ctg   768
Leu Ala Asp Ser Gly Ile Pro Gly Asn Ser Phe Asn Val Ser Ser Leu
                245                 250                 255

gtt gag ctg gat ctg tcc tat aac aag ctt aaa aac ata cca act gtc   816
Val Glu Leu Asp Leu Ser Tyr Asn Lys Leu Lys Asn Ile Pro Thr Val
            260                 265                 270

aat gaa aac ctt gaa aac tat tac ctg gag gtc aat caa ctt gag aag   864
Asn Glu Asn Leu Glu Asn Tyr Tyr Leu Glu Val Asn Gln Leu Glu Lys
        275                 280                 285

ttt gac ata aag agc ttc tgc aag atc ctg ggg cca tta tcc tac tcc   912
Phe Asp Ile Lys Ser Phe Cys Lys Ile Leu Gly Pro Leu Ser Tyr Ser
```

```
    290                 295                 300

aag atc aag cat ttg cgt ttg gat ggc aat cgc atc tca gaa acc agt   960
Lys Ile Lys His Leu Arg Leu Asp Gly Asn Arg Ile Ser Glu Thr Ser
305                 310                 315                 320

ctt cca ccg gat atg tat gaa tgt cta cgt gtt gct aac gaa gtc act   1008
Leu Pro Pro Asp Met Tyr Glu Cys Leu Arg Val Ala Asn Glu Val Thr
                325                 330                 335

ctt aat taa                                                      1017
Leu Asn  *


<210> 4
<211> 338
<212> PRT
<213> human

<400> 4
Met Ser Leu Ser Ala Phe Thr Leu Phe Leu Ala Leu Ile Gly Gly Thr
 1               5                  10                  15
Ser Gly Gln Tyr Tyr Asp Tyr Asp Phe Pro Leu Ser Ile Tyr Gly Gln
            20                  25                  30
Ser Ser Pro Asn Cys Ala Pro Glu Cys Asn Cys Pro Glu Ser Tyr Pro
        35                  40                  45
Ser Ala Met Tyr Cys Asp Glu Leu Lys Leu Lys Ser Val Pro Met Val
    50                  55                  60
Pro Pro Gly Ile Lys Tyr Leu Tyr Leu Arg Asn Asn Gln Ile Asp His
65                  70                  75                  80
Ile Asp Glu Lys Ala Phe Glu Asn Val Thr Asp Leu Gln Trp Leu Ile
                85                  90                  95
Leu Asp His Asn Leu Leu Glu Asn Ser Lys Ile Lys Gly Arg Val Phe
            100                 105                 110
Ser Lys Leu Lys Gln Leu Lys Lys Leu His Ile Asn His Asn Asn Leu
        115                 120                 125
Thr Glu Ser Val Gly Pro Leu Pro Lys Ser Leu Glu Asp Leu Gln Leu
    130                 135                 140
Thr His Asn Lys Ile Thr Lys Leu Gly Ser Phe Glu Gly Leu Val Asn
145                 150                 155                 160
Leu Thr Phe Ile His Leu Gln His Asn Arg Leu Lys Glu Asp Ala Val
                165                 170                 175
Ser Ala Ala Phe Lys Gly Leu Lys Ser Leu Glu Tyr Leu Asp Leu Ser
            180                 185                 190
Phe Asn Gln Ile Ala Arg Leu Pro Ser Gly Leu Pro Val Ser Leu Leu
        195                 200                 205
Thr Leu Tyr Leu Asp Asn Asn Lys Ile Ser Asn Ile Pro Asp Glu Tyr
    210                 215                 220
Phe Lys Arg Phe Asn Ala Leu Gln Tyr Leu Arg Leu Ser His Asn Glu
225                 230                 235                 240
Leu Ala Asp Ser Gly Ile Pro Gly Asn Ser Phe Asn Val Ser Ser Leu
                245                 250                 255
Val Glu Leu Asp Leu Ser Tyr Asn Lys Leu Lys Asn Ile Pro Thr Val
            260                 265                 270
Asn Glu Asn Leu Glu Asn Tyr Tyr Leu Glu Val Asn Gln Leu Glu Lys
        275                 280                 285
Phe Asp Ile Lys Ser Phe Cys Lys Ile Leu Gly Pro Leu Ser Tyr Ser
    290                 295                 300
Lys Ile Lys His Leu Arg Leu Asp Gly Asn Arg Ile Ser Glu Thr Ser
305                 310                 315                 320
Leu Pro Pro Asp Met Tyr Glu Cys Leu Arg Val Ala Asn Glu Val Thr
```

```
                325                 330                 335
Leu Asn



<210> 5
<211> 441
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(441)

<400> 5
atg gtg atg ggc ctg ggc gtt ttg ttg ttg gtc ttc gtg ctg ggt ctg   48
Met Val Met Gly Leu Gly Val Leu Leu Leu Val Phe Val Leu Gly Leu
 1               5                  10                  15

ggt ctg acc cca ccg acc ctg gct cag gat aac tcc agg tac aca cac   96
Gly Leu Thr Pro Pro Thr Leu Ala Gln Asp Asn Ser Arg Tyr Thr His
            20                  25                  30

ttc ctg acc cag cac tat gat gcc aaa cca cag ggc cgg gat gac aga   144
Phe Leu Thr Gln His Tyr Asp Ala Lys Pro Gln Gly Arg Asp Asp Arg
        35                  40                  45

tac tgt gaa agc atc atg agg aga cgg ggc ctg acc tca ccc tgc aaa   192
Tyr Cys Glu Ser Ile Met Arg Arg Arg Gly Leu Thr Ser Pro Cys Lys
    50                  55                  60

gac atc aac aca ttt att cat ggc aac aag cgc agc atc aag gcc atc   240
Asp Ile Asn Thr Phe Ile His Gly Asn Lys Arg Ser Ile Lys Ala Ile
65                  70                  75                  80

tgt gaa aac gag aat gga aac cct cac aga gaa aac cta aga ata agc   288
Cys Glu Asn Glu Asn Gly Asn Pro His Arg Glu Asn Leu Arg Ile Ser
                85                  90                  95

aag tct tct ttc cag gtc acc act tgc aag cta cat gga ggg tcc ccc   336
Lys Ser Ser Phe Gln Val Thr Thr Cys Lys Leu His Gly Gly Ser Pro
            100                 105                 110

tgg cct cca tgc cag tac cga gcc aca gcg ggg ttc aga aac gtt gtt   384
Trp Pro Pro Cys Gln Tyr Arg Ala Thr Ala Gly Phe Arg Asn Val Val
        115                 120                 125

gtt gct tgt gaa aat ggc tta cct gtc cac ttg gat cag tca att ttc   432
Val Ala Cys Glu Asn Gly Leu Pro Val His Leu Asp Gln Ser Ile Phe
    130                 135                 140

cgt cgt ccg                                                      441
Arg Arg Pro
145


<210> 6
<211> 147
<212> PRT
<213> human
```

<400> 6
Met Val Met Gly Leu Gly Val Leu Leu Leu Val Phe Val Leu Gly Leu
1 5 10 15
Gly Leu Thr Pro Pro Thr Leu Ala Gln Asp Asn Ser Arg Tyr Thr His
20 25 30
Phe Leu Thr Gln His Tyr Asp Ala Lys Pro Gln Gly Arg Asp Asp Arg
35 40 45
Tyr Cys Glu Ser Ile Met Arg Arg Arg Gly Leu Thr Ser Pro Cys Lys
50 55 60
Asp Ile Asn Thr Phe Ile His Gly Asn Lys Arg Ser Ile Lys Ala Ile
65 70 75 80
Cys Glu Asn Glu Asn Gly Asn Pro His Arg Glu Asn Leu Arg Ile Ser
85 90 95
Lys Ser Ser Phe Gln Val Thr Thr Cys Lys Leu His Gly Gly Ser Pro
100 105 110
Trp Pro Pro Cys Gln Tyr Arg Ala Thr Ala Gly Phe Arg Asn Val Val
115 120 125
Val Ala Cys Glu Asn Gly Leu Pro Val His Leu Asp Gln Ser Ile Phe
130 135 140
Arg Arg Pro
145

<210> 7
<211> 444
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(444)

<400> 7
atg gct ctg cag agg acc cat tca ttg ctt ctg ctt ttg ctg ctg acc 48
Met Ala Leu Gln Arg Thr His Ser Leu Leu Leu Leu Leu Leu Leu Thr
1 5 10 15

ctg ctg ggg ctg ggg ctg gtc cag ccc tcc tat ggc cag gat ggc atg 96
Leu Leu Gly Leu Gly Leu Val Gln Pro Ser Tyr Gly Gln Asp Gly Met
20 25 30

tac cag cga ttc ctg cgg caa cac gtg cac cct gag gag aca ggt ggc 144
Tyr Gln Arg Phe Leu Arg Gln His Val His Pro Glu Glu Thr Gly Gly
35 40 45

agt gat cgc tac tgc aac ttg atg atg caa aga cgg aag atg act ttg 192
Ser Asp Arg Tyr Cys Asn Leu Met Met Gln Arg Arg Lys Met Thr Leu
50 55 60

tat cac tgc aag cgc ttc aac acc ttc atc cat gaa gat atc tgg aac 240
Tyr His Cys Lys Arg Phe Asn Thr Phe Ile His Glu Asp Ile Trp Asn
65 70 75 80

att cgt agt atc tgc agc acc acc aat atc caa tgc aag aac ggc aag 288
Ile Arg Ser Ile Cys Ser Thr Thr Asn Ile Gln Cys Lys Asn Gly Lys
85 90 95

atg aac tgc cat gag ggt gta gtg aag gtc aca gat tgc agg gac aca 336
Met Asn Cys His Glu Gly Val Val Lys Val Thr Asp Cys Arg Asp Thr
100 105 110

```
gga agt tcc agg gca ccc aac tgc aga tat cgg gcc ata gcg agc act   384
Gly Ser Ser Arg Ala Pro Asn Cys Arg Tyr Arg Ala Ile Ala Ser Thr
        115                 120                 125

aga cgt gtt gtc att gcc tgt gag ggt aac cca cag gtg cct gtg cac   432
Arg Arg Val Val Ile Ala Cys Glu Gly Asn Pro Gln Val Pro Val His
    130                 135                 140

ttt gac ggt tag                                                  444
Phe Asp Gly  *
145
```

<210> 8
<211> 147
<212> PRT
<213> human

<400> 8

```
Met Ala Leu Gln Arg Thr His Ser Leu Leu Leu Leu Leu Leu Leu Thr
 1               5                  10                  15
Leu Leu Gly Leu Gly Leu Val Gln Pro Ser Tyr Gly Gln Asp Gly Met
            20                  25                  30
Tyr Gln Arg Phe Leu Arg Gln His Val His Pro Glu Glu Thr Gly Gly
        35                  40                  45
Ser Asp Arg Tyr Cys Asn Leu Met Met Gln Arg Arg Lys Met Thr Leu
    50                  55                  60
Tyr His Cys Lys Arg Phe Asn Thr Phe Ile His Glu Asp Ile Trp Asn
65                  70                  75                  80
Ile Arg Ser Ile Cys Ser Thr Thr Asn Ile Gln Cys Lys Asn Gly Lys
                85                  90                  95
Met Asn Cys His Glu Gly Val Val Lys Val Thr Asp Cys Arg Asp Thr
            100                 105                 110
Gly Ser Ser Arg Ala Pro Asn Cys Arg Tyr Arg Ala Ile Ala Ser Thr
        115                 120                 125
Arg Arg Val Val Ile Ala Cys Glu Gly Asn Pro Gln Val Pro Val His
    130                 135                 140
Phe Asp Gly
145
```

<210> 9
<211> 315
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(315)

<400> 9

```
atg agc tcc gca gcc agg tcc cgc ctc acc cgc gcc acc cgc cag gag   48
Met Ser Ser Ala Ala Arg Ser Arg Leu Thr Arg Ala Thr Arg Gln Glu
 1               5                  10                  15

atg ctg ttc ttg gcg ttg ctg ctc ctg cca gtt gtg gtc gcc ttc gcc   96
Met Leu Phe Leu Ala Leu Leu Leu Leu Pro Val Val Val Ala Phe Ala
            20                  25                  30

aga gct gaa gct gaa gaa gat ggg gac ctg cag tgc ctg tgt gtg aag   144
Arg Ala Glu Ala Glu Glu Asp Gly Asp Leu Gln Cys Leu Cys Val Lys
```

```
         35                  40                  45

acc acc tcc cag gtc cgt ccc agg cac atc acc agc ctg gag gtg atc   192
Thr Thr Ser Gln Val Arg Pro Arg His Ile Thr Ser Leu Glu Val Ile
     50                  55                  60

aag gcc gga ccc cac tgc ccc act gcc caa ctc ata gcc acg ctg aag   240
Lys Ala Gly Pro His Cys Pro Thr Ala Gln Leu Ile Ala Thr Leu Lys
 65                  70                  75                  80

aat ggg agg aaa att tgc ttg gat ctg caa gcc ctg ctg tac aag aaa   288
Asn Gly Arg Lys Ile Cys Leu Asp Leu Gln Ala Leu Leu Tyr Lys Lys
                 85                  90                  95

atc att aag gaa cat ttg gag agt tag                              315
Ile Ile Lys Glu His Leu Glu Ser  *
            100


<210> 10
<211> 104
<212> PRT
<213> human

<400> 10
Met Ser Ser Ala Ala Arg Ser Arg Leu Thr Arg Ala Thr Arg Gln Glu
 1               5                  10                  15
Met Leu Phe Leu Ala Leu Leu Leu Leu Pro Val Val Val Ala Phe Ala
            20                  25                  30
Arg Ala Glu Ala Glu Glu Asp Gly Asp Leu Gln Cys Leu Cys Val Lys
        35                  40                  45
Thr Thr Ser Gln Val Arg Pro Arg His Ile Thr Ser Leu Glu Val Ile
    50                  55                  60
Lys Ala Gly Pro His Cys Pro Thr Ala Gln Leu Ile Ala Thr Leu Lys
65                  70                  75                  80
Asn Gly Arg Lys Ile Cys Leu Asp Leu Gln Ala Leu Leu Tyr Lys Lys
                85                  90                  95
Ile Ile Lys Glu His Leu Glu Ser
            100


<210> 11
<211> 333
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(333)

<400> 11
atg agg ttc atg act ctc ctc ttc ctg aca gct ctg gca gga gcc ctg   48
Met Arg Phe Met Thr Leu Leu Phe Leu Thr Ala Leu Ala Gly Ala Leu
 1               5                  10                  15

gtc tgt gcc tat gat cca gag gcc gcc tct gcc cca gga tcg ggg aac   96
Val Cys Ala Tyr Asp Pro Glu Ala Ala Ser Ala Pro Gly Ser Gly Asn
            20                  25                  30

cct tgc cat gaa gca tca gca gct caa aag gaa aat gca ggt gaa gac   144
Pro Cys His Glu Ala Ser Ala Ala Gln Lys Glu Asn Ala Gly Glu Asp
```

```
         35                  40                  45

cca ggg tta gcc aga cag gca cca aag cca agg aag cag aga tcc agc   192
Pro Gly Leu Ala Arg Gln Ala Pro Lys Pro Arg Lys Gln Arg Ser Ser
     50                  55                  60

ctt ctg gaa aaa ggc cta gac gga gca aaa aaa gct gtg ggg gga ctc   240
Leu Leu Glu Lys Gly Leu Asp Gly Ala Lys Lys Ala Val Gly Gly Leu
 65                  70                  75                  80

gga aaa cta gga aaa gat gca gtc gaa gat cta gaa agc gtg ggt aaa   288
Gly Lys Leu Gly Lys Asp Ala Val Glu Asp Leu Glu Ser Val Gly Lys
                85                  90                  95

gga gcc gtc cat gac gtt aaa gac gtc ctt gac tca gta cta tag       333
Gly Ala Val His Asp Val Lys Asp Val Leu Asp Ser Val Leu  *
            100                 105                 110



<210> 12
<211> 110
<212> PRT
<213> human

<400> 12
Met Arg Phe Met Thr Leu Leu Phe Leu Thr Ala Leu Ala Gly Ala Leu
 1               5                  10                  15
Val Cys Ala Tyr Asp Pro Glu Ala Ala Ser Ala Pro Gly Ser Gly Asn
            20                  25                  30
Pro Cys His Glu Ala Ser Ala Ala Gln Lys Glu Asn Ala Gly Glu Asp
        35                  40                  45
Pro Gly Leu Ala Arg Gln Ala Pro Lys Pro Arg Lys Gln Arg Ser Ser
    50                  55                  60
Leu Leu Glu Lys Gly Leu Asp Gly Ala Lys Lys Ala Val Gly Gly Leu
65                  70                  75                  80
Gly Lys Leu Gly Lys Asp Ala Val Glu Asp Leu Glu Ser Val Gly Lys
                85                  90                  95
Gly Ala Val His Asp Val Lys Asp Val Leu Asp Ser Val Leu
            100                 105                 110


<210> 13
<211> 942
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(942)

<400> 13
atg gag ctg gac aga gct gtg ggg gtc ctg ggc gct gcc acc ctg ctg   48
Met Glu Leu Asp Arg Ala Val Gly Val Leu Gly Ala Ala Thr Leu Leu
 1               5                  10                  15

ctc tct ttc ctg ggc atg gcc tgg gct ctc cag gcg gca gac acc tgt   96
Leu Ser Phe Leu Gly Met Ala Trp Ala Leu Gln Ala Ala Asp Thr Cys
            20                  25                  30

cca gag gtg aag atg gtg ggc ctg gag ggc tct gac aag ctc acc att   144
```

```
Pro Glu Val Lys Met Val Gly Leu Glu Gly Ser Asp Lys Leu Thr Ile
         35                  40                  45

ctc cga ggc tgt ccg ggg ctg cct ggg gcc cct ggc gac aag gga gag   192
Leu Arg Gly Cys Pro Gly Leu Pro Gly Ala Pro Gly Asp Lys Gly Glu
     50                  55                  60

gca ggc acc aat gga aag aga gga gaa cgt ggc ccc cct gga cct cct   240
Ala Gly Thr Asn Gly Lys Arg Gly Glu Arg Gly Pro Pro Gly Pro Pro
 65                  70                  75                  80

ggg aag gca gga cca cct ggg ccc aac gga gca cct ggg gag ccc cag   288
Gly Lys Ala Gly Pro Pro Gly Pro Asn Gly Ala Pro Gly Glu Pro Gln
                 85                  90                  95

ccg tgc ctg aca ggc ccg cgt acc tgc aag gac ctg cta gac cga ggg   336
Pro Cys Leu Thr Gly Pro Arg Thr Cys Lys Asp Leu Leu Asp Arg Gly
            100                 105                 110

cac ttc ctg agc ggc tgg cac acc atc tac ctg ccc gac tgc cgg ccc   384
His Phe Leu Ser Gly Trp His Thr Ile Tyr Leu Pro Asp Cys Arg Pro
        115                 120                 125

ctg act gtg ctc tgt gac atg gac acg gac gga ggg ggc tgg acc gtt   432
Leu Thr Val Leu Cys Asp Met Asp Thr Asp Gly Gly Gly Trp Thr Val
    130                 135                 140

ttc cag cgg agg gtg gat ggc tct gtg gac ttc tac cgg gac tgg gcc   480
Phe Gln Arg Arg Val Asp Gly Ser Val Asp Phe Tyr Arg Asp Trp Ala
145                 150                 155                 160

acg tac aag cag ggc ttc ggc agt cgg ctg ggg gag ttc tgg ctg ggg   528
Thr Tyr Lys Gln Gly Phe Gly Ser Arg Leu Gly Glu Phe Trp Leu Gly
                165                 170                 175

aat gac aac atc cac gcc ctg acc gcc cag gga acc agc gag ctc cgt   576
Asn Asp Asn Ile His Ala Leu Thr Ala Gln Gly Thr Ser Glu Leu Arg
            180                 185                 190

gta gac ctg gtg gac ttt gag gac aac tac cag ttt gct aag tac aga   624
Val Asp Leu Val Asp Phe Glu Asp Asn Tyr Gln Phe Ala Lys Tyr Arg
        195                 200                 205

tca ttc aag gtg gcc gac gag gcg gag aag tac aat ctg gtc ctg ggg   672
Ser Phe Lys Val Ala Asp Glu Ala Glu Lys Tyr Asn Leu Val Leu Gly
    210                 215                 220

gcc ttc gtg gag ggc agt gcg gga gat tcc ctg acg ttc cac aac aac   720
Ala Phe Val Glu Gly Ser Ala Gly Asp Ser Leu Thr Phe His Asn Asn
225                 230                 235                 240

cag tcc ttc tcc acc aaa gac cag gac aat gat ctt aac acc gga aat   768
Gln Ser Phe Ser Thr Lys Asp Gln Asp Asn Asp Leu Asn Thr Gly Asn
                245                 250                 255

tgt gct gtg atg ttt cag gga gct tgg tgg tac aaa aac tgc cat gtg   816
Cys Ala Val Met Phe Gln Gly Ala Trp Trp Tyr Lys Asn Cys His Val
            260                 265                 270

tca aac ctg aat ggt cgc tac ctc agg ggg act cat ggc agc ttt gca   864
Ser Asn Leu Asn Gly Arg Tyr Leu Arg Gly Thr His Gly Ser Phe Ala
```

```
        275                 280                 285

aat ggc atc aac tgg aag tcg ggg aaa gga tac aat tat agc tac aag   912
Asn Gly Ile Asn Trp Lys Ser Gly Lys Gly Tyr Asn Tyr Ser Tyr Lys
    290                 295                 300

gtg tca gag atg aag gtg cga cct gcc tag                           942
Val Ser Glu Met Lys Val Arg Pro Ala  *
305                 310
```

<210> 14
<211> 313
<212> PRT
<213> human

<400> 14

```
Met Glu Leu Asp Arg Ala Val Gly Val Leu Gly Ala Ala Thr Leu Leu
 1               5                  10                  15
Leu Ser Phe Leu Gly Met Ala Trp Ala Leu Gln Ala Ala Asp Thr Cys
            20                  25                  30
Pro Glu Val Lys Met Val Gly Leu Glu Gly Ser Asp Lys Leu Thr Ile
        35                  40                  45
Leu Arg Gly Cys Pro Gly Leu Pro Gly Ala Pro Gly Asp Lys Gly Glu
    50                  55                  60
Ala Gly Thr Asn Gly Lys Arg Gly Glu Arg Gly Pro Pro Gly Pro Pro
65                  70                  75                  80
Gly Lys Ala Gly Pro Pro Gly Pro Asn Gly Ala Pro Gly Glu Pro Gln
                85                  90                  95
Pro Cys Leu Thr Gly Pro Arg Thr Cys Lys Asp Leu Leu Asp Arg Gly
            100                 105                 110
His Phe Leu Ser Gly Trp His Thr Ile Tyr Leu Pro Asp Cys Arg Pro
        115                 120                 125
Leu Thr Val Leu Cys Asp Met Asp Thr Asp Gly Gly Gly Trp Thr Val
    130                 135                 140
Phe Gln Arg Arg Val Asp Gly Ser Val Asp Phe Tyr Arg Asp Trp Ala
145                 150                 155                 160
Thr Tyr Lys Gln Gly Phe Gly Ser Arg Leu Gly Glu Phe Trp Leu Gly
                165                 170                 175
Asn Asp Asn Ile His Ala Leu Thr Ala Gln Gly Thr Ser Glu Leu Arg
            180                 185                 190
Val Asp Leu Val Asp Phe Glu Asp Asn Tyr Gln Phe Ala Lys Tyr Arg
        195                 200                 205
Ser Phe Lys Val Ala Asp Glu Ala Glu Lys Tyr Asn Leu Val Leu Gly
    210                 215                 220
Ala Phe Val Glu Gly Ser Ala Gly Asp Ser Leu Thr Phe His Asn Asn
225                 230                 235                 240
Gln Ser Phe Ser Thr Lys Asp Gln Asp Asn Asp Leu Asn Thr Gly Asn
                245                 250                 255
Cys Ala Val Met Phe Gln Gly Ala Trp Trp Tyr Lys Asn Cys His Val
            260                 265                 270
Ser Asn Leu Asn Gly Arg Tyr Leu Arg Gly Thr His Gly Ser Phe Ala
        275                 280                 285
Asn Gly Ile Asn Trp Lys Ser Gly Lys Gly Tyr Asn Tyr Ser Tyr Lys
    290                 295                 300
Val Ser Glu Met Lys Val Arg Pro Ala
305                 310
```

<210> 15
<211> 819

<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(819)

<400> 15

```
atg gtg ttg ctc acc gcg gtc ctc ctg ctg ctg gcc gcc tat gcg ggg   48
Met Val Leu Leu Thr Ala Val Leu Leu Leu Ala Ala Tyr Ala Gly
 1               5                  10                  15

ccg gcc cag agc ctg ggc tcc ttc gtg cac tgc gag ccc tgc gac gag   96
Pro Ala Gln Ser Leu Gly Ser Phe Val His Cys Glu Pro Cys Asp Glu
            20                  25                  30

aaa gcc ctc tcc atg tgc ccc ccc agc ccc ctg ggc tgc gag ctg gtc  144
Lys Ala Leu Ser Met Cys Pro Pro Ser Pro Leu Gly Cys Glu Leu Val
        35                  40                  45

aag gag ccg ggc tgc ggc tgc tgc atg acc tgc gcc ctg gcc gag ggg  192
Lys Glu Pro Gly Cys Gly Cys Cys Met Thr Cys Ala Leu Ala Glu Gly
    50                  55                  60

cag tcg tgc ggc gtc tac acc gag cgc tgc gcc cag ggg ctg cgc tgc  240
Gln Ser Cys Gly Val Tyr Thr Glu Arg Cys Ala Gln Gly Leu Arg Cys
65                  70                  75                  80

ctc ccc cgg cag gac gag gag aag ccg ctg cac gcc ctg ctg cac ggc  288
Leu Pro Arg Gln Asp Glu Glu Lys Pro Leu His Ala Leu Leu His Gly
                85                  90                  95

cgc ggg gtt tgc ctc aac gaa aag agc tac cgc gag caa gtc aag atc  336
Arg Gly Val Cys Leu Asn Glu Lys Ser Tyr Arg Glu Gln Val Lys Ile
            100                 105                 110

gag aga gac tcc cgt gag cac gag gag ccc acc acc tct gag atg gcc  384
Glu Arg Asp Ser Arg Glu His Glu Glu Pro Thr Thr Ser Glu Met Ala
        115                 120                 125

gag gag acc tac tcc ccc aag atc ttc cgg ccc aaa cac acc cgc atc  432
Glu Glu Thr Tyr Ser Pro Lys Ile Phe Arg Pro Lys His Thr Arg Ile
    130                 135                 140

tcc gag ctg aag gct gaa gca gtg aag aag gac cgc aga aag aag ctg  480
Ser Glu Leu Lys Ala Glu Ala Val Lys Lys Asp Arg Arg Lys Lys Leu
145                 150                 155                 160

acc cag tcc aag ttt gtc ggg gga gcc gag aac act gcc cac ccc cgg  528
Thr Gln Ser Lys Phe Val Gly Gly Ala Glu Asn Thr Ala His Pro Arg
                165                 170                 175

atc atc tct gca cct gag atg aga cag gag tct gag cag ggc ccc tgc  576
Ile Ile Ser Ala Pro Glu Met Arg Gln Glu Ser Glu Gln Gly Pro Cys
            180                 185                 190

cgc aga cac atg gag gct tcc ctg cag gag ctc aaa gcc agc cca cgc  624
Arg Arg His Met Glu Ala Ser Leu Gln Glu Leu Lys Ala Ser Pro Arg
        195                 200                 205

atg gtg ccc cgt gct gtg tac ctg ccc aat tgt gac cgc aaa gga ttc  672
```

```
Met Val Pro Arg Ala Val Tyr Leu Pro Asn Cys Asp Arg Lys Gly Phe
    210                 215                 220

tac aag aga aag cag tgc aaa cct tcc cgt ggc cgc aaa cgt ggc atc   720
Tyr Lys Arg Lys Gln Cys Lys Pro Ser Arg Gly Arg Lys Arg Gly Ile
225                 230                 235                 240

tgc tgg tgc gtg gac aag tac ggg atg aag ctg cca ggc atg gag tac   768
Cys Trp Cys Val Asp Lys Tyr Gly Met Lys Leu Pro Gly Met Glu Tyr
                245                 250                 255

gtt gac ggg gac ttt cag tgc cac acc ttc gac agc agc aac gtt gag   816
Val Asp Gly Asp Phe Gln Cys His Thr Phe Asp Ser Ser Asn Val Glu
            260                 265                 270

tga                                                               819
 *
```

<210> 16
<211> 272
<212> PRT
<213> human

<400> 16

```
Met Val Leu Leu Thr Ala Val Leu Leu Leu Leu Ala Ala Tyr Ala Gly
 1               5                  10                  15
Pro Ala Gln Ser Leu Gly Ser Phe Val His Cys Glu Pro Cys Asp Glu
            20                  25                  30
Lys Ala Leu Ser Met Cys Pro Pro Ser Pro Leu Gly Cys Glu Leu Val
        35                  40                  45
Lys Glu Pro Gly Cys Gly Cys Cys Met Thr Cys Ala Leu Ala Glu Gly
    50                  55                  60
Gln Ser Cys Gly Val Tyr Thr Glu Arg Cys Ala Gln Gly Leu Arg Cys
65                  70                  75                  80
Leu Pro Arg Gln Asp Glu Glu Lys Pro Leu His Ala Leu Leu His Gly
                85                  90                  95
Arg Gly Val Cys Leu Asn Glu Lys Ser Tyr Arg Glu Gln Val Lys Ile
            100                 105                 110
Glu Arg Asp Ser Arg Glu His Glu Glu Pro Thr Thr Ser Glu Met Ala
        115                 120                 125
Glu Glu Thr Tyr Ser Pro Lys Ile Phe Arg Pro Lys His Thr Arg Ile
    130                 135                 140
Ser Glu Leu Lys Ala Glu Ala Val Lys Lys Asp Arg Arg Lys Lys Leu
145                 150                 155                 160
Thr Gln Ser Lys Phe Val Gly Gly Ala Glu Asn Thr Ala His Pro Arg
                165                 170                 175
Ile Ile Ser Ala Pro Glu Met Arg Gln Glu Ser Glu Gln Gly Pro Cys
            180                 185                 190
Arg Arg His Met Glu Ala Ser Leu Gln Glu Leu Lys Ala Ser Pro Arg
        195                 200                 205
Met Val Pro Arg Ala Val Tyr Leu Pro Asn Cys Asp Arg Lys Gly Phe
    210                 215                 220
Tyr Lys Arg Lys Gln Cys Lys Pro Ser Arg Gly Arg Lys Arg Gly Ile
225                 230                 235                 240
Cys Trp Cys Val Asp Lys Tyr Gly Met Lys Leu Pro Gly Met Glu Tyr
                245                 250                 255
Val Asp Gly Asp Phe Gln Cys His Thr Phe Asp Ser Ser Asn Val Glu
            260                 265                 270
```

<210> 17
<211> 477
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(477)

<400> 17
```
atg atg cag aag cta ctc aaa tgc agt cgg ctt gtc ctg gct ctt gcc   48
Met Met Gln Lys Leu Leu Lys Cys Ser Arg Leu Val Leu Ala Leu Ala
 1               5                  10                  15

ctc atc ctg gtt ctg gaa tcc tca gtt caa ggt tat cct acg cag aga   96
Leu Ile Leu Val Leu Glu Ser Ser Val Gln Gly Tyr Pro Thr Gln Arg
            20                  25                  30

gcc agg tac caa tgg gtg cgc tgc aat cca gac agt aat tct gca aac   144
Ala Arg Tyr Gln Trp Val Arg Cys Asn Pro Asp Ser Asn Ser Ala Asn
        35                  40                  45

tgc ctt gaa gaa aaa gga cca atg ttc gaa cta ctt cca ggt gaa tcc   192
Cys Leu Glu Glu Lys Gly Pro Met Phe Glu Leu Leu Pro Gly Glu Ser
    50                  55                  60

aac aag atc ccc cgt ctg agg act gac ctt ttt cca aag acg aga atc   240
Asn Lys Ile Pro Arg Leu Arg Thr Asp Leu Phe Pro Lys Thr Arg Ile
65                  70                  75                  80

cag gac ttg aat cgt atc ttc cca ctt tct gag gac tac tct gga tca   288
Gln Asp Leu Asn Arg Ile Phe Pro Leu Ser Glu Asp Tyr Ser Gly Ser
                85                  90                  95

ggc ttc ggc tcc ggc tcc ggc tct gga tca gga tct ggg agt ggc ttc   336
Gly Phe Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Phe
            100                 105                 110

cta acg gaa atg gaa cag gat tac caa cta gta gac gaa agt gat gct   384
Leu Thr Glu Met Glu Gln Asp Tyr Gln Leu Val Asp Glu Ser Asp Ala
        115                 120                 125

ttc cat gac aac ctt agg tct ctt gac agg aat ctg ccc tca gac agc   432
Phe His Asp Asn Leu Arg Ser Leu Asp Arg Asn Leu Pro Ser Asp Ser
    130                 135                 140

cag gac ttg ggt caa cat gga tta gaa gag gat ttt atg tta taa       477
Gln Asp Leu Gly Gln His Gly Leu Glu Glu Asp Phe Met Leu  *
145                 150                 155
```

<210> 18
<211> 158
<212> PRT
<213> human

<400> 18
```
Met Met Gln Lys Leu Leu Lys Cys Ser Arg Leu Val Leu Ala Leu Ala
 1               5                  10                  15
```

```
Leu Ile Leu Val Leu Glu Ser Ser Val Gln Gly Tyr Pro Thr Gln Arg
            20                  25                  30
Ala Arg Tyr Gln Trp Val Arg Cys Asn Pro Asp Ser Asn Ser Ala Asn
        35                  40                  45
Cys Leu Glu Glu Lys Gly Pro Met Phe Glu Leu Leu Pro Gly Glu Ser
    50                  55                  60
Asn Lys Ile Pro Arg Leu Arg Thr Asp Leu Phe Pro Lys Thr Arg Ile
65                  70                  75                  80
Gln Asp Leu Asn Arg Ile Phe Pro Leu Ser Glu Asp Tyr Ser Gly Ser
                85                  90                  95
Gly Phe Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Phe
            100                 105                 110
Leu Thr Glu Met Glu Gln Asp Tyr Gln Leu Val Asp Glu Ser Asp Ala
        115                 120                 125
Phe His Asp Asn Leu Arg Ser Leu Asp Arg Asn Leu Pro Ser Asp Ser
    130                 135                 140
Gln Asp Leu Gly Gln His Gly Leu Glu Glu Asp Phe Met Leu
145                 150                 155


<210> 19
<211> 282
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(282)

<400> 19
atg aag atc tcc gtg gct gcc att ccc ttc ttc ctc ctc atc acc atc   48
Met Lys Ile Ser Val Ala Ala Ile Pro Phe Phe Leu Leu Ile Thr Ile
 1               5                  10                  15

gcc cta ggg acc aag act gaa tcc tcc tca cgg gga cct tac cac ccc   96
Ala Leu Gly Thr Lys Thr Glu Ser Ser Ser Arg Gly Pro Tyr His Pro
            20                  25                  30

tca gag tgc tgc ttc acc tac act acc tac aag atc ccg cgt cag cgg   144
Ser Glu Cys Cys Phe Thr Tyr Thr Thr Tyr Lys Ile Pro Arg Gln Arg
        35                  40                  45

att atg gat tac tat gag acc aac agc cag tgc tcc aag ccc gga att   192
Ile Met Asp Tyr Tyr Glu Thr Asn Ser Gln Cys Ser Lys Pro Gly Ile
    50                  55                  60

gtc ttc atc acc aaa agg ggc cat tcc gtc tgt acc aac ccc agt gac   240
Val Phe Ile Thr Lys Arg Gly His Ser Val Cys Thr Asn Pro Ser Asp
65                  70                  75                  80

aag tgg gtc cag gac tat atc aag gac atg aag gag aac tga           282
Lys Trp Val Gln Asp Tyr Ile Lys Asp Met Lys Glu Asn  *
                85                  90


<210> 20
<211> 93
<212> PRT
<213> human
```

<400> 20
Met Lys Ile Ser Val Ala Ala Ile Pro Phe Phe Leu Leu Ile Thr Ile
1 5 10 15
Ala Leu Gly Thr Lys Thr Glu Ser Ser Ser Arg Gly Pro Tyr His Pro
20 25 30
Ser Glu Cys Cys Phe Thr Tyr Thr Thr Tyr Lys Ile Pro Arg Gln Arg
35 40 45
Ile Met Asp Tyr Tyr Glu Thr Asn Ser Gln Cys Ser Lys Pro Gly Ile
50 55 60
Val Phe Ile Thr Lys Arg Gly His Ser Val Cys Thr Asn Pro Ser Asp
65 70 75 80
Lys Trp Val Gln Asp Tyr Ile Lys Asp Met Lys Glu Asn
85 90

<210> 21
<211> 363
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(363)

<400> 21
atg aag gtc tcc gag gct gcc ctg tct ctc ctt gtc ctc atc ctt atc 48
Met Lys Val Ser Glu Ala Ala Leu Ser Leu Leu Val Leu Ile Leu Ile
1 5 10 15

att act tcg gct tct cgc agc cag cca aaa gtt cct gag tgg gtg aac 96
Ile Thr Ser Ala Ser Arg Ser Gln Pro Lys Val Pro Glu Trp Val Asn
20 25 30

acc cca tcc acc tgc tgc ctg aag tat tat gag aaa gtg ttg cca agg 144
Thr Pro Ser Thr Cys Cys Leu Lys Tyr Tyr Glu Lys Val Leu Pro Arg
35 40 45

aga cta gtg gtg gga tac aga aag gcc ctc aac tgt cac ctg cca gca 192
Arg Leu Val Val Gly Tyr Arg Lys Ala Leu Asn Cys His Leu Pro Ala
50 55 60

atc atc ttc gtc acc aag agg aac cga gaa gtc tgc acc aac ccc aat 240
Ile Ile Phe Val Thr Lys Arg Asn Arg Glu Val Cys Thr Asn Pro Asn
65 70 75 80

gac gac tgg gtc caa gag tac atc aag gat ccc aac cta cct ttg ctg 288
Asp Asp Trp Val Gln Glu Tyr Ile Lys Asp Pro Asn Leu Pro Leu Leu
85 90 95

cct acc agg aac ttg tcc acg gtt aaa att att aca gca aag aat ggt 336
Pro Thr Arg Asn Leu Ser Thr Val Lys Ile Ile Thr Ala Lys Asn Gly
100 105 110

caa ccc cag ctc ctc aac tcc cag tga 363
Gln Pro Gln Leu Leu Asn Ser Gln *
115 120

<210> 22
<211> 120
<212> PRT

<213> human

<400> 22

```
Met Lys Val Ser Glu Ala Ala Leu Ser Leu Leu Val Leu Ile Leu Ile
 1               5                  10                  15
Ile Thr Ser Ala Ser Arg Ser Gln Pro Lys Val Pro Glu Trp Val Asn
            20                  25                  30
Thr Pro Ser Thr Cys Cys Leu Lys Tyr Tyr Glu Lys Val Leu Pro Arg
        35                  40                  45
Arg Leu Val Val Gly Tyr Arg Lys Ala Leu Asn Cys His Leu Pro Ala
    50                  55                  60
Ile Ile Phe Val Thr Lys Arg Asn Arg Glu Val Cys Thr Asn Pro Asn
65                  70                  75                  80
Asp Asp Trp Val Gln Glu Tyr Ile Lys Asp Pro Asn Leu Pro Leu Leu
                85                  90                  95
Pro Thr Arg Asn Leu Ser Thr Val Lys Ile Ile Thr Ala Lys Asn Gly
            100                 105                 110
Gln Pro Gln Leu Leu Asn Ser Gln
        115                 120
```

<210> 23
<211> 270
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(270)

<400> 23

```
atg aag ggc ctt gca gct gcc ctc ctt gtc ctc gtc tgc acc atg gcc   48
Met Lys Gly Leu Ala Ala Ala Leu Leu Val Leu Val Cys Thr Met Ala
 1               5                  10                  15

ctc tgc tcc tgt gca caa gtt ggt acc aac aaa gag ctc tgc tgc ctc   96
Leu Cys Ser Cys Ala Gln Val Gly Thr Asn Lys Glu Leu Cys Cys Leu
            20                  25                  30

gtc tat acc tcc tgg cag att cca caa aag ttc ata gtt gac tat tct   144
Val Tyr Thr Ser Trp Gln Ile Pro Gln Lys Phe Ile Val Asp Tyr Ser
        35                  40                  45

gaa acc agc ccc cag tgc ccc aag cca ggt gtc atc ctc cta acc aag   192
Glu Thr Ser Pro Gln Cys Pro Lys Pro Gly Val Ile Leu Leu Thr Lys
    50                  55                  60

aga ggc cgg cag atc tgt gct gac ccc aat aag aag tgg gtc cag aaa   240
Arg Gly Arg Gln Ile Cys Ala Asp Pro Asn Lys Lys Trp Val Gln Lys
65                  70                  75                  80

tac atc agc gac ctg aag ctg aat gcc tga                          270
Tyr Ile Ser Asp Leu Lys Leu Asn Ala  *
                85
```

<210> 24
<211> 89
<212> PRT
<213> human

<400> 24
Met Lys Gly Leu Ala Ala Ala Leu Leu Val Leu Val Cys Thr Met Ala
1 5 10 15
Leu Cys Ser Cys Ala Gln Val Gly Thr Asn Lys Glu Leu Cys Cys Leu
20 25 30
Val Tyr Thr Ser Trp Gln Ile Pro Gln Lys Phe Ile Val Asp Tyr Ser
35 40 45
Glu Thr Ser Pro Gln Cys Pro Lys Pro Gly Val Ile Leu Leu Thr Lys
50 55 60
Arg Gly Arg Gln Ile Cys Ala Asp Pro Asn Lys Lys Trp Val Gln Lys
65 70 75 80
Tyr Ile Ser Asp Leu Lys Leu Asn Ala
85

<210> 25
<211> 1194
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(1194)

<400> 25
atg gct ccg ata tct ctg tcg tgg ctg ctc cgc ttg gcc acc ttc tgc 48
Met Ala Pro Ile Ser Leu Ser Trp Leu Leu Arg Leu Ala Thr Phe Cys
1 5 10 15

cat ctg act gtc ctg ctg gct gga cag cac cac ggt gtg acg aaa tgc 96
His Leu Thr Val Leu Leu Ala Gly Gln His His Gly Val Thr Lys Cys
20 25 30

aac atc acg tgc agc aag atg aca tca aag ata cct gta gct ttg ctc 144
Asn Ile Thr Cys Ser Lys Met Thr Ser Lys Ile Pro Val Ala Leu Leu
35 40 45

atc cac tat caa cag aac cag gca tca tgc ggc aaa cgc gca atc atc 192
Ile His Tyr Gln Gln Asn Gln Ala Ser Cys Gly Lys Arg Ala Ile Ile
50 55 60

ttg gag acg aga cag cac agg ctg ttc tgt gcc gac ccg aag gag caa 240
Leu Glu Thr Arg Gln His Arg Leu Phe Cys Ala Asp Pro Lys Glu Gln
65 70 75 80

tgg gtc aag gac gcg atg cag cat ctg gac cgc cag gct gct gcc cta 288
Trp Val Lys Asp Ala Met Gln His Leu Asp Arg Gln Ala Ala Ala Leu
85 90 95

act cga aat ggc ggc acc ttc gag aag cag atc ggc gag gtg aag ccc 336
Thr Arg Asn Gly Gly Thr Phe Glu Lys Gln Ile Gly Glu Val Lys Pro
100 105 110

agg acc acc cct gcc gcc ggg gga atg gac gag tct gtg gtc ctg gag 384
Arg Thr Thr Pro Ala Ala Gly Gly Met Asp Glu Ser Val Val Leu Glu
115 120 125

ccc gaa gcc aca ggc gaa agc agt agc ctg gag ccg act cct tct tcc 432
Pro Glu Ala Thr Gly Glu Ser Ser Ser Leu Glu Pro Thr Pro Ser Ser
130 135 140

```
cag gaa gca cag agg gcc ctg ggg acc tcc cca gag ctg ccg acg ggc   480
Gln Glu Ala Gln Arg Ala Leu Gly Thr Ser Pro Glu Leu Pro Thr Gly
145                 150                 155                 160

gtg act ggt tcc tca ggg acc agg ctc ccc ccg acg cca aag gct cag   528
Val Thr Gly Ser Ser Gly Thr Arg Leu Pro Pro Thr Pro Lys Ala Gln
                165                 170                 175

gat gga ggg cct gtg ggc acg gag ctt ttc cga gtg cct ccc gtc tcc   576
Asp Gly Gly Pro Val Gly Thr Glu Leu Phe Arg Val Pro Pro Val Ser
            180                 185                 190

act gcc gcc acg tgg cag agt tct gct ccc cac caa cct ggg ccc agc   624
Thr Ala Ala Thr Trp Gln Ser Ser Ala Pro His Gln Pro Gly Pro Ser
        195                 200                 205

ctc tgg gct gag gca aag acc tct gag gcc ccg tcc acc cag gac ccc   672
Leu Trp Ala Glu Ala Lys Thr Ser Glu Ala Pro Ser Thr Gln Asp Pro
    210                 215                 220

tcc acc cag gcc tcc act gcg tcc tcc cca gcc cca gag gag aat gct   720
Ser Thr Gln Ala Ser Thr Ala Ser Ser Pro Ala Pro Glu Glu Asn Ala
225                 230                 235                 240

ccg tct gaa ggc cag cgt gtg tgg ggt cag gga cag agc ccc agg cca   768
Pro Ser Glu Gly Gln Arg Val Trp Gly Gln Gly Gln Ser Pro Arg Pro
                245                 250                 255

gag aac tct ctg gag cgg gag gag atg ggt ccc gtg cca gcg cac acg   816
Glu Asn Ser Leu Glu Arg Glu Glu Met Gly Pro Val Pro Ala His Thr
            260                 265                 270

gat gcc ttc cag gac tgg ggg cct ggc agc atg gcc cac gtc tct gtg   864
Asp Ala Phe Gln Asp Trp Gly Pro Gly Ser Met Ala His Val Ser Val
        275                 280                 285

gtc cct gtc tcc tca gaa ggg acc ccc agc agg gag cca gtg gct tca   912
Val Pro Val Ser Ser Glu Gly Thr Pro Ser Arg Glu Pro Val Ala Ser
    290                 295                 300

ggc agc tgg acc cct aag gct gag gaa ccc atc cat gcc acc atg gac   960
Gly Ser Trp Thr Pro Lys Ala Glu Glu Pro Ile His Ala Thr Met Asp
305                 310                 315                 320

ccc cag agg ctg ggc gtc ctt atc act cct gtc cct gac gcc cag gct  1008
Pro Gln Arg Leu Gly Val Leu Ile Thr Pro Val Pro Asp Ala Gln Ala
                325                 330                 335

gcc acc cgg agg cag gcg gtg ggg ctg ctg gcc ttc ctt ggc ctc ctc  1056
Ala Thr Arg Arg Gln Ala Val Gly Leu Leu Ala Phe Leu Gly Leu Leu
            340                 345                 350

ttc tgc ctg ggg gtg gcc atg ttc acc tac cag agc ctc cag ggc tgc  1104
Phe Cys Leu Gly Val Ala Met Phe Thr Tyr Gln Ser Leu Gln Gly Cys
        355                 360                 365

cct cga aag atg gca gga gag atg gcg gag ggc ctt cgc tac atc ccc  1152
Pro Arg Lys Met Ala Gly Glu Met Ala Glu Gly Leu Arg Tyr Ile Pro
    370                 375                 380

cgg agc tgt ggt agt aat tca tat gtc ctg gtg ccc gtg tga          1194
```

```
Arg Ser Cys Gly Ser Asn Ser Tyr Val Leu Val Pro Val  *
385                 390                 395



<210> 26
<211> 397
<212> PRT
<213> human

<400> 26
Met Ala Pro Ile Ser Leu Ser Trp Leu Leu Arg Leu Ala Thr Phe Cys
 1               5                  10                  15
His Leu Thr Val Leu Leu Ala Gly Gln His His Gly Val Thr Lys Cys
            20                  25                  30
Asn Ile Thr Cys Ser Lys Met Thr Ser Lys Ile Pro Val Ala Leu Leu
        35                  40                  45
Ile His Tyr Gln Gln Asn Gln Ala Ser Cys Gly Lys Arg Ala Ile Ile
    50                  55                  60
Leu Glu Thr Arg Gln His Arg Leu Phe Cys Ala Asp Pro Lys Glu Gln
65                  70                  75                  80
Trp Val Lys Asp Ala Met Gln His Leu Asp Arg Gln Ala Ala Ala Leu
                85                  90                  95
Thr Arg Asn Gly Gly Thr Phe Glu Lys Gln Ile Gly Glu Val Lys Pro
            100                 105                 110
Arg Thr Thr Pro Ala Ala Gly Gly Met Asp Glu Ser Val Val Leu Glu
        115                 120                 125
Pro Glu Ala Thr Gly Glu Ser Ser Ser Leu Glu Pro Thr Pro Ser Ser
    130                 135                 140
Gln Glu Ala Gln Arg Ala Leu Gly Thr Ser Pro Glu Leu Pro Thr Gly
145                 150                 155                 160
Val Thr Gly Ser Ser Gly Thr Arg Leu Pro Pro Thr Pro Lys Ala Gln
                165                 170                 175
Asp Gly Gly Pro Val Gly Thr Glu Leu Phe Arg Val Pro Pro Val Ser
            180                 185                 190
Thr Ala Ala Thr Trp Gln Ser Ser Ala Pro His Gln Pro Gly Pro Ser
        195                 200                 205
Leu Trp Ala Glu Ala Lys Thr Ser Glu Ala Pro Ser Thr Gln Asp Pro
    210                 215                 220
Ser Thr Gln Ala Ser Thr Ala Ser Ser Pro Ala Pro Glu Glu Asn Ala
225                 230                 235                 240
Pro Ser Glu Gly Gln Arg Val Trp Gly Gln Gly Gln Ser Pro Arg Pro
                245                 250                 255
Glu Asn Ser Leu Glu Arg Glu Glu Met Gly Pro Val Pro Ala His Thr
            260                 265                 270
Asp Ala Phe Gln Asp Trp Gly Pro Gly Ser Met Ala His Val Ser Val
        275                 280                 285
Val Pro Val Ser Ser Glu Gly Thr Pro Ser Arg Glu Pro Val Ala Ser
    290                 295                 300
Gly Ser Trp Thr Pro Lys Ala Glu Glu Pro Ile His Ala Thr Met Asp
305                 310                 315                 320
Pro Gln Arg Leu Gly Val Leu Ile Thr Pro Val Pro Asp Ala Gln Ala
                325                 330                 335
Ala Thr Arg Arg Gln Ala Val Gly Leu Leu Ala Phe Leu Gly Leu Leu
            340                 345                 350
Phe Cys Leu Gly Val Ala Met Phe Thr Tyr Gln Ser Leu Gln Gly Cys
        355                 360                 365
Pro Arg Lys Met Ala Gly Glu Met Ala Glu Gly Leu Arg Tyr Ile Pro
    370                 375                 380
Arg Ser Cys Gly Ser Asn Ser Tyr Val Leu Val Pro Val
385                 390                 395
```

<210> 27
<211> 282
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(282)

<400> 27
```
atg aac gcc aag gtc gtg gtc gtg ctg gtc ctc gtg ctg acc gcg ctc   48
Met Asn Ala Lys Val Val Val Val Leu Val Leu Val Leu Thr Ala Leu
 1               5                  10                  15

tgc ctc agc gac ggg aag ccc gtc agc ctg agc tac aga tgc cca tgc   96
Cys Leu Ser Asp Gly Lys Pro Val Ser Leu Ser Tyr Arg Cys Pro Cys
            20                  25                  30

cga ttc ttc gaa agc cat gtt gcc aga gcc aac gtc aag cat ctc aaa   144
Arg Phe Phe Glu Ser His Val Ala Arg Ala Asn Val Lys His Leu Lys
        35                  40                  45

att ctc aac act cca aac tgt gcc ctt cag att gta gcc cgg ctg aag   192
Ile Leu Asn Thr Pro Asn Cys Ala Leu Gln Ile Val Ala Arg Leu Lys
    50                  55                  60

aac aac aac aga caa gtg tgc att gac ccg aag cta aag tgg att cag   240
Asn Asn Asn Arg Gln Val Cys Ile Asp Pro Lys Leu Lys Trp Ile Gln
65                  70                  75                  80

gag tac ctg gag aaa gct tta aac aag agg ttc aag atg tga           282
Glu Tyr Leu Glu Lys Ala Leu Asn Lys Arg Phe Lys Met  *
                85                  90
```

<210> 28
<211> 93
<212> PRT
<213> human

<400> 28
```
Met Asn Ala Lys Val Val Val Val Leu Val Leu Val Leu Thr Ala Leu
 1               5                  10                  15
Cys Leu Ser Asp Gly Lys Pro Val Ser Leu Ser Tyr Arg Cys Pro Cys
            20                  25                  30
Arg Phe Phe Glu Ser His Val Ala Arg Ala Asn Val Lys His Leu Lys
        35                  40                  45
Ile Leu Asn Thr Pro Asn Cys Ala Leu Gln Ile Val Ala Arg Leu Lys
    50                  55                  60
Asn Asn Asn Arg Gln Val Cys Ile Asp Pro Lys Leu Lys Trp Ile Gln
65                  70                  75                  80
Glu Tyr Leu Glu Lys Ala Leu Asn Lys Arg Phe Lys Met
                85                  90
```

<210> 29
<211> 4551
<212> DNA

<213> human

<220>
<221> CDS
<222> (1)...(4551)

<400> 29

```
atg gct ccc tac ccc tgt ggc tgc cac atc ctg ctg ctg ctc ttc tgc   48
Met Ala Pro Tyr Pro Cys Gly Cys His Ile Leu Leu Leu Leu Phe Cys
 1               5                  10                  15

tgc ctg gcg gct gcc cgg gcc aac ctg ctg aac ctg aac tgg ctt tgg   96
Cys Leu Ala Ala Ala Arg Ala Asn Leu Leu Asn Leu Asn Trp Leu Trp
            20                  25                  30

ttc aat aat gag gac acc agc cac gca gct acc acg atc cct gag ccc   144
Phe Asn Asn Glu Asp Thr Ser His Ala Ala Thr Thr Ile Pro Glu Pro
        35                  40                  45

cag ggg ccc ctg cct gtg cag ccc aca gca gat acc acc aca cac gtg   192
Gln Gly Pro Leu Pro Val Gln Pro Thr Ala Asp Thr Thr Thr His Val
    50              55                  60

acc ccc cgg aat ggt tcc aca gag cca gcg aca gcc cct ggc agc cct   240
Thr Pro Arg Asn Gly Ser Thr Glu Pro Ala Thr Ala Pro Gly Ser Pro
65                  70                  75                  80

gag cca ccc tca gag ctg ctg gaa gat ggc cag gac acc ccc act tct   288
Glu Pro Pro Ser Glu Leu Leu Glu Asp Gly Gln Asp Thr Pro Thr Ser
                85                  90                  95

gcc gag agc ccg gac gcg cca gag gag aac att gcc ggt gtc gga gcc   336
Ala Glu Ser Pro Asp Ala Pro Glu Glu Asn Ile Ala Gly Val Gly Ala
            100                 105                 110

gag atc ctg aac gtg gcc aaa ggc atc cgg agc ttc gtc cag ctg tgg   384
Glu Ile Leu Asn Val Ala Lys Gly Ile Arg Ser Phe Val Gln Leu Trp
        115                 120                 125

aat gac act gtc ccc act gag agc ttg gcc agg gcg gaa acc ctg gtc   432
Asn Asp Thr Val Pro Thr Glu Ser Leu Ala Arg Ala Glu Thr Leu Val
    130                 135                 140

ctg gag act cct gtg ggc ccc ctt gcc ctc gct ggg cct tcc agc acc   480
Leu Glu Thr Pro Val Gly Pro Leu Ala Leu Ala Gly Pro Ser Ser Thr
145                 150                 155                 160

ccc cag gag aat ggg acc act ctc tgg ccc agc cgt ggc att cct agc   528
Pro Gln Glu Asn Gly Thr Thr Leu Trp Pro Ser Arg Gly Ile Pro Ser
                165                 170                 175

tct ccg ggc gcc cac aca acc gag gct ggc acc ttg cct gca ccc acc   576
Ser Pro Gly Ala His Thr Thr Glu Ala Gly Thr Leu Pro Ala Pro Thr
            180                 185                 190

cca tcg cct ccg tcc ctg ggc agg ccc tgg gca cca ctc acg ggg ccc   624
Pro Ser Pro Pro Ser Leu Gly Arg Pro Trp Ala Pro Leu Thr Gly Pro
        195                 200                 205

tca gtg cca cca cca tct tca gag cgc atc agc gag gag gtg ggg ctg   672
Ser Val Pro Pro Pro Ser Ser Glu Arg Ile Ser Glu Glu Val Gly Leu
```

```
    210                 215                 220

ctg cag ctc ctt ggg gac ccc ccg ccc cag cag gtc acc cag acg gat   720
Leu Gln Leu Leu Gly Asp Pro Pro Pro Gln Gln Val Thr Gln Thr Asp
225                 230                 235                 240

gac ccc gac gtc ggg ctg gcc tac gtc ttt ggg cca gat gcc aac agt   768
Asp Pro Asp Val Gly Leu Ala Tyr Val Phe Gly Pro Asp Ala Asn Ser
                245                 250                 255

ggc caa gtg gcc cgg tac cac ttc ccc agc ctc ttc ttc cgt gac ttc   816
Gly Gln Val Ala Arg Tyr His Phe Pro Ser Leu Phe Phe Arg Asp Phe
            260                 265                 270

tca ctg ctg ttc cac atc cgg cca gcc aca gag ggc cca ggg gtg ctg   864
Ser Leu Leu Phe His Ile Arg Pro Ala Thr Glu Gly Pro Gly Val Leu
        275                 280                 285

ttc gcc atc acg gac tcg gcg cag gcc atg gtc ttg ctg ggc gtg aag   912
Phe Ala Ile Thr Asp Ser Ala Gln Ala Met Val Leu Leu Gly Val Lys
    290                 295                 300

ctc tct ggg gtg cag gac ggg cac cag gac atc tcc ctg ctc tac aca   960
Leu Ser Gly Val Gln Asp Gly His Gln Asp Ile Ser Leu Leu Tyr Thr
305                 310                 315                 320

gaa cct ggt gca ggc cag acc cac aca gcc gcc agc ttc cgg ctc ccc  1008
Glu Pro Gly Ala Gly Gln Thr His Thr Ala Ala Ser Phe Arg Leu Pro
                325                 330                 335

gcc ttc gtc ggc cag tgg aca cac tta gcc ctc agt gtg gca ggt ggc  1056
Ala Phe Val Gly Gln Trp Thr His Leu Ala Leu Ser Val Ala Gly Gly
            340                 345                 350

ttt gtg gcc ctc tac gtg gac tgt gag gag ttc cag aga atg ccg ctt  1104
Phe Val Ala Leu Tyr Val Asp Cys Glu Glu Phe Gln Arg Met Pro Leu
        355                 360                 365

gct cgg tcc tca cgg ggc ctg gag ctg gag cct ggc gcc ggg ctc ttc  1152
Ala Arg Ser Ser Arg Gly Leu Glu Leu Glu Pro Gly Ala Gly Leu Phe
    370                 375                 380

gtg gct cag gcg ggg gga gcg gac cct gac aag ttc cag ggg gtg atc  1200
Val Ala Gln Ala Gly Gly Ala Asp Pro Asp Lys Phe Gln Gly Val Ile
385                 390                 395                 400

gct gag ctg aag gtg cgc agg gac ccc cag gtg agc ccc atg cac tgc  1248
Ala Glu Leu Lys Val Arg Arg Asp Pro Gln Val Ser Pro Met His Cys
                405                 410                 415

ctg gac gag gaa ggc gat gac tca gat ggg gca ttc gga gac tct ggc  1296
Leu Asp Glu Glu Gly Asp Asp Ser Asp Gly Ala Phe Gly Asp Ser Gly
            420                 425                 430

agc ggg ctc ggg gac gcc cgg gag ctt ctc agg gag gag acg ggc gcg  1344
Ser Gly Leu Gly Asp Ala Arg Glu Leu Leu Arg Glu Glu Thr Gly Ala
        435                 440                 445

gcc cta aaa ccc agg ctc ccc gcg cca ccc ccc gtc acc acg cca ccc  1392
Ala Leu Lys Pro Arg Leu Pro Ala Pro Pro Pro Val Thr Thr Pro Pro
    450                 455                 460
```

```
ttg gct gga ggc agc agc acg gaa gat tcc aga agt gaa gaa gtc gag   1440
Leu Ala Gly Gly Ser Ser Thr Glu Asp Ser Arg Ser Glu Glu Val Glu
465                 470                 475                 480

gag cag acc acg gtg gct tcg tta gga gct cag aca ctt cct ggc tca   1488
Glu Gln Thr Thr Val Ala Ser Leu Gly Ala Gln Thr Leu Pro Gly Ser
                485                 490                 495

gat tct gtc tcc acg tgg gac ggg agt gtc cgg acc cct ggg ggc cgc   1536
Asp Ser Val Ser Thr Trp Asp Gly Ser Val Arg Thr Pro Gly Gly Arg
            500                 505                 510

gtg aaa gag ggc ggc ctg aag ggg cag aaa ggg gag cca ggt gtt ccg   1584
Val Lys Glu Gly Gly Leu Lys Gly Gln Lys Gly Glu Pro Gly Val Pro
        515                 520                 525

ggc cca cct ggc cgg gca ggc ccc cca gga tcc cca tgc cta cct ggt   1632
Gly Pro Pro Gly Arg Ala Gly Pro Pro Gly Ser Pro Cys Leu Pro Gly
    530                 535                 540

ccc ccg ggt ctc ccg tgc cca gtg agt ccc ctg ggt cct gca ggc cca   1680
Pro Pro Gly Leu Pro Cys Pro Val Ser Pro Leu Gly Pro Ala Gly Pro
545                 550                 555                 560

gcg ttg caa act gtc ccc gga cca caa gga ccc cca ggg cct ccg ggg   1728
Ala Leu Gln Thr Val Pro Gly Pro Gln Gly Pro Pro Gly Pro Pro Gly
                565                 570                 575

agg gac ggc acc cct gga agg gac ggc gag ccg ggc gac ccc ggt gaa   1776
Arg Asp Gly Thr Pro Gly Arg Asp Gly Glu Pro Gly Asp Pro Gly Glu
            580                 585                 590

gac gga aag ccg ggc gac acc ggg cca caa ggc ttc cct ggg act cca   1824
Asp Gly Lys Pro Gly Asp Thr Gly Pro Gln Gly Phe Pro Gly Thr Pro
        595                 600                 605

ggg gat gta ggt ccc aag gga gac aag gga gac cct ggg gtt gga gag   1872
Gly Asp Val Gly Pro Lys Gly Asp Lys Gly Asp Pro Gly Val Gly Glu
    610                 615                 620

aga ggg ccc cca gga ccc caa ggg cct cca ggg ccc cca gga ccc tcc   1920
Arg Gly Pro Pro Gly Pro Gln Gly Pro Pro Gly Pro Pro Gly Pro Ser
625                 630                 635                 640

ttc aga cac gac aag ctg acc ttc att gac atg gag gga tct ggc ttt   1968
Phe Arg His Asp Lys Leu Thr Phe Ile Asp Met Glu Gly Ser Gly Phe
                645                 650                 655

ggg ggc gat ctg gag gcc ctg cgg ggt cct cga ggc ttc cct gga cct   2016
Gly Gly Asp Leu Glu Ala Leu Arg Gly Pro Arg Gly Phe Pro Gly Pro
            660                 665                 670

ccc gga ccc ccc ggt gtc cca ggc ctg ccc ggc gag cca ggc cgc ttt   2064
Pro Gly Pro Pro Gly Val Pro Gly Leu Pro Gly Glu Pro Gly Arg Phe
        675                 680                 685

ggg gtg aac agc tcc gac gtc cca gga ccc gcc ggc ctt cct ggt gtg   2112
Gly Val Asn Ser Ser Asp Val Pro Gly Pro Ala Gly Leu Pro Gly Val
    690                 695                 700
```

```
cct ggg cgc gag ggt ccc ccc ggg ttt cct ggc ctc ccg gga ccc cca   2160
Pro Gly Arg Glu Gly Pro Pro Gly Phe Pro Gly Leu Pro Gly Pro Pro
705                 710                 715                 720

ggc cct ccg gga aga gag ggg ccc cca gga agg act ggg cag aaa ggc   2208
Gly Pro Pro Gly Arg Glu Gly Pro Pro Gly Arg Thr Gly Gln Lys Gly
                725                 730                 735

agc ctg ggt gaa gca ggc gcc cca gga cat aag ggg agc aag gga gcc   2256
Ser Leu Gly Glu Ala Gly Ala Pro Gly His Lys Gly Ser Lys Gly Ala
            740                 745                 750

ccc ggt cct gct ggt gct cgt ggg gag agc ggc ctg gca gga gcc ccc   2304
Pro Gly Pro Ala Gly Ala Arg Gly Glu Ser Gly Leu Ala Gly Ala Pro
        755                 760                 765

gga cct gct gga cca cca ggc ccc cct ggg ccc cct ggg ccc cca gga   2352
Gly Pro Ala Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly
    770                 775                 780

cca gga ctc ccc gct gga ttt gat gac atg gaa ggc tcc ggg ggg ccc   2400
Pro Gly Leu Pro Ala Gly Phe Asp Asp Met Glu Gly Ser Gly Gly Pro
785                 790                 795                 800

ttc tgg tca aca gcc cga agc gct gat ggg cca cag gga cct ccc ggc   2448
Phe Trp Ser Thr Ala Arg Ser Ala Asp Gly Pro Gln Gly Pro Pro Gly
                805                 810                 815

ctg ccg gga ctt aag ggg gat cct ggc gtg cct ggg ctg ccg ggg gcg   2496
Leu Pro Gly Leu Lys Gly Asp Pro Gly Val Pro Gly Leu Pro Gly Ala
            820                 825                 830

aag gga gaa gtt gga gca gat gga atc ccc ggg ttc ccc ggc ctc cct   2544
Lys Gly Glu Val Gly Ala Asp Gly Ile Pro Gly Phe Pro Gly Leu Pro
        835                 840                 845

ggc aga gag ggc att gct ggg ccc cag ggg cca aag gga gac aga ggc   2592
Gly Arg Glu Gly Ile Ala Gly Pro Gln Gly Pro Lys Gly Asp Arg Gly
    850                 855                 860

agc cgg gga gaa aag gga gat cca ggg aag gac gga gtc ggg cag ccg   2640
Ser Arg Gly Glu Lys Gly Asp Pro Gly Lys Asp Gly Val Gly Gln Pro
865                 870                 875                 880

ggc ctc cct ggc ccc ccc gga ccc ccg gga cct gtg gtc tac gtg tcg   2688
Gly Leu Pro Gly Pro Pro Gly Pro Pro Gly Pro Val Val Tyr Val Ser
                885                 890                 895

gag cag gac gga tcc gtc ctg agc gtg ccg gga cct gag ggc cgg ccg   2736
Glu Gln Asp Gly Ser Val Leu Ser Val Pro Gly Pro Glu Gly Arg Pro
            900                 905                 910

ggt ttc gca ggc ttt ccc gga cct gca gga ccc aag ggc aac ctg ggc   2784
Gly Phe Ala Gly Phe Pro Gly Pro Ala Gly Pro Lys Gly Asn Leu Gly
        915                 920                 925

tct aag ggc gaa cga ggc tcc ccg gga ccc aag ggt gag aag ggt gaa   2832
Ser Lys Gly Glu Arg Gly Ser Pro Gly Pro Lys Gly Glu Lys Gly Glu
    930                 935                 940

ccg ggc agc atc ttc agc ccc gac ggc ggt gcc ctg ggc cct gcc cag   2880
```

```
Pro Gly Ser Ile Phe Ser Pro Asp Gly Gly Ala Leu Gly Pro Ala Gln
945                 950                 955                 960

aaa gga gcc aag gga gag ccg ggc ttc cga gga ccc ccg ggt cca tac   2928
Lys Gly Ala Lys Gly Glu Pro Gly Phe Arg Gly Pro Pro Gly Pro Tyr
                965                 970                 975

gga cgg ccg ggg tac aag gga gag att ggc ttt cct gga cgg ccg ggt   2976
Gly Arg Pro Gly Tyr Lys Gly Glu Ile Gly Phe Pro Gly Arg Pro Gly
            980                 985                 990

cgc ccc ggg atg aac gga ttg aaa gga gag aaa ggg gag ccg gga gat   3024
Arg Pro Gly Met Asn Gly Leu Lys Gly Glu Lys Gly Glu Pro Gly Asp
        995                 1000                1005

gcc agc ctt gga ttt ggc atg agg gga atg ccc ggc ccc cca gga cct   3072
Ala Ser Leu Gly Phe Gly Met Arg Gly Met Pro Gly Pro Pro Gly Pro
    1010                1015                1020

cca ggg ccc cca ggc cct cca ggg act cct gtt tac gac agc aat gtg   3120
Pro Gly Pro Pro Gly Pro Pro Gly Thr Pro Val Tyr Asp Ser Asn Val
1025                1030                1035                1040

ttt gct gag tcc agc cgc ccc ggg cct cca gga ttg cca ggg aat cag   3168
Phe Ala Glu Ser Ser Arg Pro Gly Pro Pro Gly Leu Pro Gly Asn Gln
                1045                1050                1055

ggc cct cca gga ccc aag ggc gcc aaa gga gaa gtg ggc ccc ccc gga   3216
Gly Pro Pro Gly Pro Lys Gly Ala Lys Gly Glu Val Gly Pro Pro Gly
            1060                1065                1070

cca cca ggg cag ttt ccg ttt gac ttt ctt cag ttg gag gct gaa atg   3264
Pro Pro Gly Gln Phe Pro Phe Asp Phe Leu Gln Leu Glu Ala Glu Met
        1075                1080                1085

aag ggg gag aag gga gac cga ggt gat gca gga cag aaa ggc gaa agg   3312
Lys Gly Glu Lys Gly Asp Arg Gly Asp Ala Gly Gln Lys Gly Glu Arg
    1090                1095                1100

ggg gag ccc ggg ggc ggc ggt ttc ttc ggc tcc agc ctg ccc ggc ccc   3360
Gly Glu Pro Gly Gly Gly Gly Phe Phe Gly Ser Ser Leu Pro Gly Pro
1105                1110                1115                1120

ccc ggc ccc cca ggc cca cgt ggc tac cct ggg att cca ggt ccc aag   3408
Pro Gly Pro Pro Gly Pro Arg Gly Tyr Pro Gly Ile Pro Gly Pro Lys
                1125                1130                1135

gga gag agc atc cgg ggc cag ccc ggc cca cct gga cct cag gga ccc   3456
Gly Glu Ser Ile Arg Gly Gln Pro Gly Pro Pro Gly Pro Gln Gly Pro
            1140                1145                1150

ccc ggc atc ggc tac gag ggg cgc cag ggc cct ccc ggc ccc cca ggc   3504
Pro Gly Ile Gly Tyr Glu Gly Arg Gln Gly Pro Pro Gly Pro Pro Gly
        1155                1160                1165

ccc cca ggg ccc cct tca ttt cct ggc cct cac agg cag act atc agc   3552
Pro Pro Gly Pro Pro Ser Phe Pro Gly Pro His Arg Gln Thr Ile Ser
    1170                1175                1180

gtt ccc ggc cct ccg ggc ccc cct ggg ccc cct ggg ccc cct gga acc   3600
Val Pro Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly Thr
```

```
1185                1190                1195                1200

atg ggc gcc tcc tca ggg gtg agg ctc tgg gct aca cgc cag gcc atg   3648
Met Gly Ala Ser Ser Gly Val Arg Leu Trp Ala Thr Arg Gln Ala Met
                1205                1210                1215

ctg ggc cag gtg cac gag gtt ccc gag ggc tgg ctc atc ttc gtg gcc   3696
Leu Gly Gln Val His Glu Val Pro Glu Gly Trp Leu Ile Phe Val Ala
            1220                1225                1230

gag cag gag gag ctc tac gtc cgc gtg cag aac ggg ttc cgg aag gtc   3744
Glu Gln Glu Glu Leu Tyr Val Arg Val Gln Asn Gly Phe Arg Lys Val
        1235                1240                1245

cag ctg gag gcc cgg aca cca ctc cca cga ggg acg gac aat gaa gtg   3792
Gln Leu Glu Ala Arg Thr Pro Leu Pro Arg Gly Thr Asp Asn Glu Val
    1250                1255                1260

gcc gcc ttg cag ccc ccc gtg gtg cag ctg cac gac agc aac ccc tac   3840
Ala Ala Leu Gln Pro Pro Val Val Gln Leu His Asp Ser Asn Pro Tyr
1265                1270                1275                1280

ccg cgg cgg gag cac ccc cac ccc acc gcg cgg ccc tgg cgg gca gat   3888
Pro Arg Arg Glu His Pro His Pro Thr Ala Arg Pro Trp Arg Ala Asp
                1285                1290                1295

gac atc ctg gcc agc ccc cct cgc ctg ccc gag ccc cag ccc tac ccc   3936
Asp Ile Leu Ala Ser Pro Pro Arg Leu Pro Glu Pro Gln Pro Tyr Pro
            1300                1305                1310

gga gcc ccg cac cac agc tcc tac gtg cac ctg cgg ccg gcg cga ccc   3984
Gly Ala Pro His His Ser Ser Tyr Val His Leu Arg Pro Ala Arg Pro
        1315                1320                1325

aca agc cca ccc gcc cac agc cac cgc gac ttc cag ccg gtg ctc cac   4032
Thr Ser Pro Pro Ala His Ser His Arg Asp Phe Gln Pro Val Leu His
    1330                1335                1340

ctg gtt gcg ctc aac agc ccc ctg tca ggc ggc atg cgg ggc atc cgc   4080
Leu Val Ala Leu Asn Ser Pro Leu Ser Gly Gly Met Arg Gly Ile Arg
1345                1350                1355                1360

ggg gcc gac ttc cag tgc ttc cag cag gcg cgg gcc gtg ggg ctg gcg   4128
Gly Ala Asp Phe Gln Cys Phe Gln Gln Ala Arg Ala Val Gly Leu Ala
                1365                1370                1375

ggc acc ttc cgc gcc ttc ctg tcc tcg cgc ctg cag gac ctg tac agc   4176
Gly Thr Phe Arg Ala Phe Leu Ser Ser Arg Leu Gln Asp Leu Tyr Ser
            1380                1385                1390

atc gtg cgc cgt gcc gac cgc gca gcc gtg ccc atc gtc aac ctc aag   4224
Ile Val Arg Arg Ala Asp Arg Ala Ala Val Pro Ile Val Asn Leu Lys
        1395                1400                1405

gac gag ctg ctg ttt ccc agc tgg gag gct ctg ttc tca ggc tct gag   4272
Asp Glu Leu Leu Phe Pro Ser Trp Glu Ala Leu Phe Ser Gly Ser Glu
    1410                1415                1420

ggt ccg ctg aag ccc ggg gca cgc atc ttc tcc ttt gac ggc aag gac   4320
Gly Pro Leu Lys Pro Gly Ala Arg Ile Phe Ser Phe Asp Gly Lys Asp
1425                1430                1435                1440
```

```
gtc ctg agg cac ccc acc tgg ccc cag aag agc gtg tgg cat ggc tcg   4368
Val Leu Arg His Pro Thr Trp Pro Gln Lys Ser Val Trp His Gly Ser
                1445                1450                1455

gac ccc aac ggg cgc agg ctg acc gag agc tac tgt gag acg tgg cgg   4416
Asp Pro Asn Gly Arg Arg Leu Thr Glu Ser Tyr Cys Glu Thr Trp Arg
            1460                1465                1470

acg gag gct ccc tcg gcc acg ggc cag gcc tcc tcg ctg ctg ggg ggc   4464
Thr Glu Ala Pro Ser Ala Thr Gly Gln Ala Ser Ser Leu Leu Gly Gly
        1475                1480                1485

agg ctc ctg ggg cag agt gcc gcg agc tgc cat cac gcc tac atc gtg   4512
Arg Leu Leu Gly Gln Ser Ala Ala Ser Cys His His Ala Tyr Ile Val
    1490                1495                1500

ctc tgc att gag aac agc ttc atg act gcc tcc aag tag               4551
Leu Cys Ile Glu Asn Ser Phe Met Thr Ala Ser Lys  *
1505                1510                1515
```

<210> 30
<211> 1516
<212> PRT
<213> human

<400> 30

```
Met Ala Pro Tyr Pro Cys Gly Cys His Ile Leu Leu Leu Leu Phe Cys
 1               5                  10                  15
Cys Leu Ala Ala Ala Arg Ala Asn Leu Leu Asn Leu Asn Trp Leu Trp
            20                  25                  30
Phe Asn Asn Glu Asp Thr Ser His Ala Ala Thr Thr Ile Pro Glu Pro
        35                  40                  45
Gln Gly Pro Leu Pro Val Gln Pro Thr Ala Asp Thr Thr Thr His Val
    50                  55                  60
Thr Pro Arg Asn Gly Ser Thr Glu Pro Ala Thr Ala Pro Gly Ser Pro
65                  70                  75                  80
Glu Pro Pro Ser Glu Leu Leu Glu Asp Gly Gln Asp Thr Pro Thr Ser
                85                  90                  95
Ala Glu Ser Pro Asp Ala Pro Glu Glu Asn Ile Ala Gly Val Gly Ala
            100                 105                 110
Glu Ile Leu Asn Val Ala Lys Gly Ile Arg Ser Phe Val Gln Leu Trp
        115                 120                 125
Asn Asp Thr Val Pro Thr Glu Ser Leu Ala Arg Ala Glu Thr Leu Val
    130                 135                 140
Leu Glu Thr Pro Val Gly Pro Leu Ala Leu Ala Gly Pro Ser Ser Thr
145                 150                 155                 160
Pro Gln Glu Asn Gly Thr Thr Leu Trp Pro Ser Arg Gly Ile Pro Ser
                165                 170                 175
Ser Pro Gly Ala His Thr Thr Glu Ala Gly Thr Leu Pro Ala Pro Thr
            180                 185                 190
Pro Ser Pro Pro Ser Leu Gly Arg Pro Trp Ala Pro Leu Thr Gly Pro
        195                 200                 205
Ser Val Pro Pro Pro Ser Ser Glu Arg Ile Ser Glu Glu Val Gly Leu
    210                 215                 220
Leu Gln Leu Leu Gly Asp Pro Pro Pro Gln Gln Val Thr Gln Thr Asp
225                 230                 235                 240
Asp Pro Asp Val Gly Leu Ala Tyr Val Phe Gly Pro Asp Ala Asn Ser
                245                 250                 255
Gly Gln Val Ala Arg Tyr His Phe Pro Ser Leu Phe Phe Arg Asp Phe
```

```
            260                 265                 270
Ser Leu Leu Phe His Ile Arg Pro Ala Thr Glu Gly Pro Gly Val Leu
        275                 280                 285
Phe Ala Ile Thr Asp Ser Ala Gln Ala Met Val Leu Leu Gly Val Lys
    290                 295                 300
Leu Ser Gly Val Gln Asp Gly His Gln Asp Ile Ser Leu Leu Tyr Thr
305                 310                 315                 320
Glu Pro Gly Ala Gly Gln Thr His Thr Ala Ala Ser Phe Arg Leu Pro
                325                 330                 335
Ala Phe Val Gly Gln Trp Thr His Leu Ala Leu Ser Val Ala Gly Gly
            340                 345                 350
Phe Val Ala Leu Tyr Val Asp Cys Glu Glu Phe Gln Arg Met Pro Leu
        355                 360                 365
Ala Arg Ser Ser Arg Gly Leu Glu Leu Glu Pro Gly Ala Gly Leu Phe
    370                 375                 380
Val Ala Gln Ala Gly Gly Ala Asp Pro Asp Lys Phe Gln Gly Val Ile
385                 390                 395                 400
Ala Glu Leu Lys Val Arg Arg Asp Pro Gln Val Ser Pro Met His Cys
                405                 410                 415
Leu Asp Glu Glu Gly Asp Asp Ser Asp Gly Ala Phe Gly Asp Ser Gly
            420                 425                 430
Ser Gly Leu Gly Asp Ala Arg Glu Leu Leu Arg Glu Glu Thr Gly Ala
        435                 440                 445
Ala Leu Lys Pro Arg Leu Pro Ala Pro Pro Pro Val Thr Thr Pro Pro
    450                 455                 460
Leu Ala Gly Gly Ser Ser Thr Glu Asp Ser Arg Ser Glu Glu Val Glu
465                 470                 475                 480
Glu Gln Thr Thr Val Ala Ser Leu Gly Ala Gln Thr Leu Pro Gly Ser
                485                 490                 495
Asp Ser Val Ser Thr Trp Asp Gly Ser Val Arg Thr Pro Gly Gly Arg
            500                 505                 510
Val Lys Glu Gly Gly Leu Lys Gly Gln Lys Gly Glu Pro Gly Val Pro
        515                 520                 525
Gly Pro Pro Gly Arg Ala Gly Pro Pro Gly Ser Pro Cys Leu Pro Gly
    530                 535                 540
Pro Pro Gly Leu Pro Cys Pro Val Ser Pro Leu Gly Pro Ala Gly Pro
545                 550                 555                 560
Ala Leu Gln Thr Val Pro Gly Pro Gln Gly Pro Pro Gly Pro Pro Gly
                565                 570                 575
Arg Asp Gly Thr Pro Gly Arg Asp Gly Glu Pro Gly Asp Pro Gly Glu
            580                 585                 590
Asp Gly Lys Pro Gly Asp Thr Gly Pro Gln Gly Phe Pro Gly Thr Pro
        595                 600                 605
Gly Asp Val Gly Pro Lys Gly Asp Lys Gly Asp Pro Gly Val Gly Glu
    610                 615                 620
Arg Gly Pro Pro Gly Pro Gln Gly Pro Pro Gly Pro Pro Gly Pro Ser
625                 630                 635                 640
Phe Arg His Asp Lys Leu Thr Phe Ile Asp Met Glu Gly Ser Gly Phe
                645                 650                 655
Gly Gly Asp Leu Glu Ala Leu Arg Gly Pro Arg Gly Phe Pro Gly Pro
            660                 665                 670
Pro Gly Pro Pro Gly Val Pro Gly Leu Pro Gly Glu Pro Gly Arg Phe
        675                 680                 685
Gly Val Asn Ser Ser Asp Val Pro Gly Pro Ala Gly Leu Pro Gly Val
    690                 695                 700
Pro Gly Arg Glu Gly Pro Pro Gly Phe Pro Gly Leu Pro Gly Pro Pro
705                 710                 715                 720
Gly Pro Pro Gly Arg Glu Gly Pro Pro Gly Arg Thr Gly Gln Lys Gly
                725                 730                 735
Ser Leu Gly Glu Ala Gly Ala Pro Gly His Lys Gly Ser Lys Gly Ala
            740                 745                 750
```

```
Pro Gly Pro Ala Gly Ala Arg Gly Glu Ser Gly Leu Ala Gly Ala Pro
        755                 760                 765
Gly Pro Ala Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly
    770                 775                 780
Pro Gly Leu Pro Ala Gly Phe Asp Asp Met Glu Gly Ser Gly Gly Pro
785                 790                 795                 800
Phe Trp Ser Thr Ala Arg Ser Ala Asp Gly Pro Gln Gly Pro Pro Gly
                805                 810                 815
Leu Pro Gly Leu Lys Gly Asp Pro Gly Val Pro Gly Leu Pro Gly Ala
            820                 825                 830
Lys Gly Glu Val Gly Ala Asp Gly Ile Pro Gly Phe Pro Gly Leu Pro
        835                 840                 845
Gly Arg Glu Gly Ile Ala Gly Pro Gln Gly Pro Lys Gly Asp Arg Gly
    850                 855                 860
Ser Arg Gly Glu Lys Gly Asp Pro Gly Lys Asp Gly Val Gly Gln Pro
865                 870                 875                 880
Gly Leu Pro Gly Pro Pro Gly Pro Pro Gly Pro Val Val Tyr Val Ser
                885                 890                 895
Glu Gln Asp Gly Ser Val Leu Ser Val Pro Gly Pro Glu Gly Arg Pro
            900                 905                 910
Gly Phe Ala Gly Phe Pro Gly Pro Ala Gly Pro Lys Gly Asn Leu Gly
        915                 920                 925
Ser Lys Gly Glu Arg Gly Ser Pro Gly Pro Lys Gly Glu Lys Gly Glu
    930                 935                 940
Pro Gly Ser Ile Phe Ser Pro Asp Gly Gly Ala Leu Gly Pro Ala Gln
945                 950                 955                 960
Lys Gly Ala Lys Gly Glu Pro Gly Phe Arg Gly Pro Pro Gly Pro Tyr
                965                 970                 975
Gly Arg Pro Gly Tyr Lys Gly Glu Ile Gly Phe Pro Gly Arg Pro Gly
            980                 985                 990
Arg Pro Gly Met Asn Gly Leu Lys Gly Glu Lys Gly Glu Pro Gly Asp
        995                 1000                1005
Ala Ser Leu Gly Phe Gly Met Arg Gly Met Pro Gly Pro Pro Gly Pro
    1010                1015                1020
Pro Gly Pro Pro Gly Pro Pro Gly Thr Pro Val Tyr Asp Ser Asn Val
1025                1030                1035                1040
Phe Ala Glu Ser Ser Arg Pro Gly Pro Pro Gly Leu Pro Gly Asn Gln
                1045                1050                1055
Gly Pro Pro Gly Pro Lys Gly Ala Lys Gly Glu Val Gly Pro Pro Gly
            1060                1065                1070
Pro Pro Gly Gln Phe Pro Phe Asp Phe Leu Gln Leu Glu Ala Glu Met
        1075                1080                1085
Lys Gly Glu Lys Gly Asp Arg Gly Asp Ala Gly Gln Lys Gly Glu Arg
    1090                1095                1100
Gly Glu Pro Gly Gly Gly Gly Phe Phe Gly Ser Ser Leu Pro Gly Pro
1105                1110                1115                1120
Pro Gly Pro Pro Gly Pro Arg Gly Tyr Pro Gly Ile Pro Gly Pro Lys
                1125                1130                1135
Gly Glu Ser Ile Arg Gly Gln Pro Gly Pro Pro Gly Pro Gln Gly Pro
            1140                1145                1150
Pro Gly Ile Gly Tyr Glu Gly Arg Gln Gly Pro Pro Gly Pro Pro Gly
        1155                1160                1165
Pro Pro Gly Pro Pro Ser Phe Pro Gly Pro His Arg Gln Thr Ile Ser
    1170                1175                1180
Val Pro Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly Thr
1185                1190                1195                1200
Met Gly Ala Ser Ser Gly Val Arg Leu Trp Ala Thr Arg Gln Ala Met
                1205                1210                1215
Leu Gly Gln Val His Glu Val Pro Glu Gly Trp Leu Ile Phe Val Ala
            1220                1225                1230
Glu Gln Glu Glu Leu Tyr Val Arg Val Gln Asn Gly Phe Arg Lys Val
```

```
        1235                1240                1245
Gln Leu Glu Ala Arg Thr Pro Leu Pro Arg Gly Thr Asp Asn Glu Val
    1250                1255                1260
Ala Ala Leu Gln Pro Pro Val Val Gln Leu His Asp Ser Asn Pro Tyr
1265                1270                1275                1280
Pro Arg Arg Glu His Pro His Pro Thr Ala Arg Pro Trp Arg Ala Asp
                1285                1290                1295
Asp Ile Leu Ala Ser Pro Pro Arg Leu Pro Glu Pro Gln Pro Tyr Pro
            1300                1305                1310
Gly Ala Pro His His Ser Ser Tyr Val His Leu Arg Pro Ala Arg Pro
        1315                1320                1325
Thr Ser Pro Pro Ala His Ser His Arg Asp Phe Gln Pro Val Leu His
    1330                1335                1340
Leu Val Ala Leu Asn Ser Pro Leu Ser Gly Gly Met Arg Gly Ile Arg
1345                1350                1355                1360
Gly Ala Asp Phe Gln Cys Phe Gln Gln Ala Arg Ala Val Gly Leu Ala
                1365                1370                1375
Gly Thr Phe Arg Ala Phe Leu Ser Ser Arg Leu Gln Asp Leu Tyr Ser
            1380                1385                1390
Ile Val Arg Arg Ala Asp Arg Ala Ala Val Pro Ile Val Asn Leu Lys
        1395                1400                1405
Asp Glu Leu Leu Phe Pro Ser Trp Glu Ala Leu Phe Ser Gly Ser Glu
    1410                1415                1420
Gly Pro Leu Lys Pro Gly Ala Arg Ile Phe Ser Phe Asp Gly Lys Asp
1425                1430                1435                1440
Val Leu Arg His Pro Thr Trp Pro Gln Lys Ser Val Trp His Gly Ser
                1445                1450                1455
Asp Pro Asn Gly Arg Arg Leu Thr Glu Ser Tyr Cys Glu Thr Trp Arg
            1460                1465                1470
Thr Glu Ala Pro Ser Ala Thr Gly Gln Ala Ser Ser Leu Leu Gly Gly
        1475                1480                1485
Arg Leu Leu Gly Gln Ser Ala Ala Ser Cys His His Ala Tyr Ile Val
    1490                1495                1500
Leu Cys Ile Glu Asn Ser Phe Met Thr Ala Ser Lys
1505                1510                1515


<210> 31
<211> 606
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(606)

<400> 31
atg gac ctc agt ctt ctc tgg gta ctt ctg ccc cta gtc acc atg gcc   48
Met Asp Leu Ser Leu Leu Trp Val Leu Leu Pro Leu Val Thr Met Ala
 1               5                  10                  15

tgg ggc cag tat ggc gat tat gga tac cca tac cag cag tat cat gac   96
Trp Gly Gln Tyr Gly Asp Tyr Gly Tyr Pro Tyr Gln Gln Tyr His Asp
            20                  25                  30

tac agc gat gat ggg tgg gtg aat ttg aac cgg caa ggc ttc agc tac   144
Tyr Ser Asp Asp Gly Trp Val Asn Leu Asn Arg Gln Gly Phe Ser Tyr
        35                  40                  45

cag tgt ccc cag ggg cag gtg ata gtg gcc gtg agg agc atc ttc agc   192
Gln Cys Pro Gln Gly Gln Val Ile Val Ala Val Arg Ser Ile Phe Ser
```

```
     50                  55                  60

aag aag gaa ggt tct gac aga caa tgg aac tac gcc tgc atg ccc acg   240
Lys Lys Glu Gly Ser Asp Arg Gln Trp Asn Tyr Ala Cys Met Pro Thr
 65                  70                  75                  80

cca cag agc ctc ggg gaa ccc acg gag tgc tgg tgg gag gag atc aac   288
Pro Gln Ser Leu Gly Glu Pro Thr Glu Cys Trp Trp Glu Glu Ile Asn
                 85                  90                  95

agg gct ggc atg gaa tgg tac cag acg tgc tcc aac aat ggg ctg gtg   336
Arg Ala Gly Met Glu Trp Tyr Gln Thr Cys Ser Asn Asn Gly Leu Val
            100                 105                 110

gca gga ttc cag agc cgc tac ttc gag tca gtg ctg gat cgg gag tgg   384
Ala Gly Phe Gln Ser Arg Tyr Phe Glu Ser Val Leu Asp Arg Glu Trp
        115                 120                 125

cag ttt tac tgt tgt cgc tac agc aag agg tgc cca tat tcc tgc tgg   432
Gln Phe Tyr Cys Cys Arg Tyr Ser Lys Arg Cys Pro Tyr Ser Cys Trp
    130                 135                 140

cta aca aca gaa tat cca ggt cac tat ggt gag gaa atg gac atg att   480
Leu Thr Thr Glu Tyr Pro Gly His Tyr Gly Glu Glu Met Asp Met Ile
145                 150                 155                 160

tcc tac aat tat gat tac tat atc cga gga gca aca acc act ttc tct   528
Ser Tyr Asn Tyr Asp Tyr Tyr Ile Arg Gly Ala Thr Thr Thr Phe Ser
                165                 170                 175

gca gtg gaa agg gat cgc cag tgg aag ttc ata atg tgc cgg atg act   576
Ala Val Glu Arg Asp Arg Gln Trp Lys Phe Ile Met Cys Arg Met Thr
            180                 185                 190

gaa tac gac tgt gaa ttt gca aat gtt tag                          606
Glu Tyr Asp Cys Glu Phe Ala Asn Val  *
        195                 200


<210> 32
<211> 201
<212> PRT
<213> human

<400> 32
Met Asp Leu Ser Leu Leu Trp Val Leu Leu Pro Leu Val Thr Met Ala
 1               5                  10                  15
Trp Gly Gln Tyr Gly Asp Tyr Gly Tyr Pro Tyr Gln Gln Tyr His Asp
            20                  25                  30
Tyr Ser Asp Asp Gly Trp Val Asn Leu Asn Arg Gln Gly Phe Ser Tyr
        35                  40                  45
Gln Cys Pro Gln Gly Gln Val Ile Val Ala Val Arg Ser Ile Phe Ser
    50                  55                  60
Lys Lys Glu Gly Ser Asp Arg Gln Trp Asn Tyr Ala Cys Met Pro Thr
65                  70                  75                  80
Pro Gln Ser Leu Gly Glu Pro Thr Glu Cys Trp Trp Glu Glu Ile Asn
                85                  90                  95
Arg Ala Gly Met Glu Trp Tyr Gln Thr Cys Ser Asn Asn Gly Leu Val
            100                 105                 110
Ala Gly Phe Gln Ser Arg Tyr Phe Glu Ser Val Leu Asp Arg Glu Trp
        115                 120                 125
```

```
Gln Phe Tyr Cys Cys Arg Tyr Ser Lys Arg Cys Pro Tyr Ser Cys Trp
    130                 135                 140
Leu Thr Thr Glu Tyr Pro Gly His Tyr Gly Glu Glu Met Asp Met Ile
145                 150                 155                 160
Ser Tyr Asn Tyr Asp Tyr Tyr Ile Arg Gly Ala Thr Thr Thr Phe Ser
                165                 170                 175
Ala Val Glu Arg Asp Arg Gln Trp Lys Phe Ile Met Cys Arg Met Thr
            180                 185                 190
Glu Tyr Asp Cys Glu Phe Ala Asn Val
        195                 200


<210> 33
<211> 369
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(369)

<400> 33
atg aag ctt ctc acg ggc ctg gtt ttc tgc tcc ttg gtc ctg agt gtc   48
Met Lys Leu Leu Thr Gly Leu Val Phe Cys Ser Leu Val Leu Ser Val
 1               5                  10                  15

agc agc cga agc ttc ttt tcg ttc ctt ggc gag gct ttt gat ggg gct   96
Ser Ser Arg Ser Phe Phe Ser Phe Leu Gly Glu Ala Phe Asp Gly Ala
            20                  25                  30

cgg gac atg tgg aga gcc tac tct gac atg aga gaa gcc aat tac atc   144
Arg Asp Met Trp Arg Ala Tyr Ser Asp Met Arg Glu Ala Asn Tyr Ile
        35                  40                  45

ggc tca gac aaa tac ttc cat gct cgg ggg aac tat gat gct gcc aaa   192
Gly Ser Asp Lys Tyr Phe His Ala Arg Gly Asn Tyr Asp Ala Ala Lys
    50                  55                  60

agg gga cct ggg ggt gcc tgg gcc gca gaa gtg atc agc aat gcc aga   240
Arg Gly Pro Gly Gly Ala Trp Ala Ala Glu Val Ile Ser Asn Ala Arg
65                  70                  75                  80

gag aat atc cag aga ctc aca ggc cat ggt gcg gag gac tcg ctg gcc   288
Glu Asn Ile Gln Arg Leu Thr Gly His Gly Ala Glu Asp Ser Leu Ala
                85                  90                  95

gat cag gct gcc aat aaa tgg ggc agg agt ggc aga gac ccc aat cac   336
Asp Gln Ala Ala Asn Lys Trp Gly Arg Ser Gly Arg Asp Pro Asn His
            100                 105                 110

ttc cga cct gct ggc ctg cct gag aaa tac tga                      369
Phe Arg Pro Ala Gly Leu Pro Glu Lys Tyr  *
        115                 120


<210> 34
<211> 122
<212> PRT
<213> human

<400> 34
```

```
Met Lys Leu Leu Thr Gly Leu Val Phe Cys Ser Leu Val Leu Ser Val
 1               5                  10                  15
Ser Ser Arg Ser Phe Phe Ser Phe Leu Gly Glu Ala Phe Asp Gly Ala
            20                  25                  30
Arg Asp Met Trp Arg Ala Tyr Ser Asp Met Arg Glu Ala Asn Tyr Ile
        35                  40                  45
Gly Ser Asp Lys Tyr Phe His Ala Arg Gly Asn Tyr Asp Ala Ala Lys
    50                  55                  60
Arg Gly Pro Gly Gly Ala Trp Ala Ala Glu Val Ile Ser Asn Ala Arg
65                  70                  75                  80
Glu Asn Ile Gln Arg Leu Thr Gly His Gly Ala Glu Asp Ser Leu Ala
                85                  90                  95
Asp Gln Ala Ala Asn Lys Trp Gly Arg Ser Gly Arg Asp Pro Asn His
            100                 105                 110
Phe Arg Pro Ala Gly Leu Pro Glu Lys Tyr
        115                 120


<210> 35
<211> 285
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(285)

<400> 35
atg ttg acc gag ctg gag aaa gcc ttg aac tct atc atc gac gtc tac   48
Met Leu Thr Glu Leu Glu Lys Ala Leu Asn Ser Ile Ile Asp Val Tyr
 1               5                  10                  15

cac aag tac tcc ctg ata aag ggg aat ttc cat gcc gtc tac agg gat   96
His Lys Tyr Ser Leu Ile Lys Gly Asn Phe His Ala Val Tyr Arg Asp
            20                  25                  30

gac ctg aag aaa ttg cta gag acc gag tgt cct cag tat atc agg aaa   144
Asp Leu Lys Lys Leu Leu Glu Thr Glu Cys Pro Gln Tyr Ile Arg Lys
        35                  40                  45

aag ggt gca gac gtc tgg ttc aaa gag ttg gat atc aac act gat ggt   192
Lys Gly Ala Asp Val Trp Phe Lys Glu Leu Asp Ile Asn Thr Asp Gly
    50                  55                  60

gca gtt aac ttc cag gag ttc ctc att ctg gtg ata aag atg ggc tgg   240
Ala Val Asn Phe Gln Glu Phe Leu Ile Leu Val Ile Lys Met Gly Trp
65                  70                  75                  80

cag ccc aca aaa aaa gcc atg aag aaa gcc aca aag agt agc tga       285
Gln Pro Thr Lys Lys Ala Met Lys Lys Ala Thr Lys Ser Ser  *
                85                  90


<210> 36
<211> 94
<212> PRT
<213> human

<400> 36
Met Leu Thr Glu Leu Glu Lys Ala Leu Asn Ser Ile Ile Asp Val Tyr
```

```
 1               5                  10                  15
His Lys Tyr Ser Leu Ile Lys Gly Asn Phe His Ala Val Tyr Arg Asp
            20                  25                  30
Asp Leu Lys Lys Leu Leu Glu Thr Glu Cys Pro Gln Tyr Ile Arg Lys
        35                  40                  45
Lys Gly Ala Asp Val Trp Phe Lys Glu Leu Asp Ile Asn Thr Asp Gly
    50                  55                  60
Ala Val Asn Phe Gln Glu Phe Leu Ile Leu Val Ile Lys Met Gly Trp
65                  70                  75                  80
Gln Pro Thr Lys Lys Ala Met Lys Lys Ala Thr Lys Ser Ser
                85                  90


<210> 37
<211> 345
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(345)

<400> 37
atg act tgc aaa atg tcg cag ctg gaa cgc aac ata gag acc atc atc   48
Met Thr Cys Lys Met Ser Gln Leu Glu Arg Asn Ile Glu Thr Ile Ile
 1               5                  10                  15

aac acc ttc cac caa tac tct gtg aag ctg ggg cac cca gac acc ctg   96
Asn Thr Phe His Gln Tyr Ser Val Lys Leu Gly His Pro Asp Thr Leu
            20                  25                  30

aac cag ggg gaa ttc aaa gag ctg gtg cga aaa gat ctg caa aat ttt   144
Asn Gln Gly Glu Phe Lys Glu Leu Val Arg Lys Asp Leu Gln Asn Phe
        35                  40                  45

ctc aag aag gag aat aag aat gaa aag gtc ata gaa cac atc atg gag   192
Leu Lys Lys Glu Asn Lys Asn Glu Lys Val Ile Glu His Ile Met Glu
    50                  55                  60

gac ctg gac aca aat gca gac aag cag ctg agc ttc gag gag ttc atc   240
Asp Leu Asp Thr Asn Ala Asp Lys Gln Leu Ser Phe Glu Glu Phe Ile
65                  70                  75                  80

atg ctg atg gcg agg cta acc tgg gcc tcc cac gag aag atg cac gag   288
Met Leu Met Ala Arg Leu Thr Trp Ala Ser His Glu Lys Met His Glu
                85                  90                  95

ggt gac gag ggc cct ggc cac cac cat aag cca ggc ctc ggg gag ggc   336
Gly Asp Glu Gly Pro Gly His His His Lys Pro Gly Leu Gly Glu Gly
            100                 105                 110

acc ccc taa                                                      345
Thr Pro  *


<210> 38
<211> 114
<212> PRT
<213> human
```

<400> 38
Met Thr Cys Lys Met Ser Gln Leu Glu Arg Asn Ile Glu Thr Ile Ile
1 5 10 15
Asn Thr Phe His Gln Tyr Ser Val Lys Leu Gly His Pro Asp Thr Leu
20 25 30
Asn Gln Gly Glu Phe Lys Glu Leu Val Arg Lys Asp Leu Gln Asn Phe
35 40 45
Leu Lys Lys Glu Asn Lys Asn Glu Lys Val Ile Glu His Ile Met Glu
50 55 60
Asp Leu Asp Thr Asn Ala Asp Lys Gln Leu Ser Phe Glu Glu Phe Ile
65 70 75 80
Met Leu Met Ala Arg Leu Thr Trp Ala Ser His Glu Lys Met His Glu
85 90 95
Gly Asp Glu Gly Pro Gly His His His Lys Pro Gly Leu Gly Glu Gly
100 105 110
Thr Pro

<210> 39
<211> 279
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(279)

<400> 39
atg aca aaa ctt gaa gag cat ctg gag gga att gtc aat atc ttc cac 48
Met Thr Lys Leu Glu Glu His Leu Glu Gly Ile Val Asn Ile Phe His
1 5 10 15

caa tac tca gtt cgg aag ggg cat ttt gac acc ctc tct aag ggt gag 96
Gln Tyr Ser Val Arg Lys Gly His Phe Asp Thr Leu Ser Lys Gly Glu
20 25 30

ctg aag cag ctg ctt aca aag gag ctt gca aac acc atc aag aat atc 144
Leu Lys Gln Leu Leu Thr Lys Glu Leu Ala Asn Thr Ile Lys Asn Ile
35 40 45

aaa gat aaa gct gtc att gat gaa ata ttc caa ggc ctg gat gct aat 192
Lys Asp Lys Ala Val Ile Asp Glu Ile Phe Gln Gly Leu Asp Ala Asn
50 55 60

caa gat gaa cag gtc gac ttt caa gaa ttc ata tcc ctg gta gcc att 240
Gln Asp Glu Gln Val Asp Phe Gln Glu Phe Ile Ser Leu Val Ala Ile
65 70 75 80

gcg ctg aag gct gcc cat tac cac acc cac aaa gag tag 279
Ala Leu Lys Ala Ala His Tyr His Thr His Lys Glu *
85 90

<210> 40
<211> 92
<212> PRT
<213> human

<400> 40
Met Thr Lys Leu Glu Glu His Leu Glu Gly Ile Val Asn Ile Phe His

```
 1               5                  10                  15
Gln Tyr Ser Val Arg Lys Gly His Phe Asp Thr Leu Ser Lys Gly Glu
            20                  25                  30
Leu Lys Gln Leu Leu Thr Lys Glu Leu Ala Asn Thr Ile Lys Asn Ile
        35                  40                  45
Lys Asp Lys Ala Val Ile Asp Glu Ile Phe Gln Gly Leu Asp Ala Asn
    50                  55                  60
Gln Asp Glu Gln Val Asp Phe Gln Glu Phe Ile Ser Leu Val Ala Ile
65                  70                  75                  80
Ala Leu Lys Ala Ala His Tyr His Thr His Lys Glu
                85                  90
```

<210> 41
<211> 903
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(903)

<400> 41

```
atg aga att gca gtg att tgc ttt tgc ctc cta ggc atc acc tgt gcc   48
Met Arg Ile Ala Val Ile Cys Phe Cys Leu Leu Gly Ile Thr Cys Ala
 1               5                  10                  15

ata cca gtt aaa cag gct gat tct gga agt tct gag gaa aag cag ctt   96
Ile Pro Val Lys Gln Ala Asp Ser Gly Ser Ser Glu Glu Lys Gln Leu
            20                  25                  30

tac aac aaa tac cca gat gct gtg gcc aca tgg cta aac cct gac cca   144
Tyr Asn Lys Tyr Pro Asp Ala Val Ala Thr Trp Leu Asn Pro Asp Pro
        35                  40                  45

tct cag aag cag aat ctc cta gcc cca cag acc ctt cca agt aag tcc   192
Ser Gln Lys Gln Asn Leu Leu Ala Pro Gln Thr Leu Pro Ser Lys Ser
    50                  55                  60

aac gaa agc cat gac cac atg gat gat atg gat gat gaa gat gat gat   240
Asn Glu Ser His Asp His Met Asp Asp Met Asp Asp Glu Asp Asp Asp
65                  70                  75                  80

gac cat gtg gac agc cag gac tcc att gac tcg aac gac tct gat gat   288
Asp His Val Asp Ser Gln Asp Ser Ile Asp Ser Asn Asp Ser Asp Asp
                85                  90                  95

gta gat gac act gat gat tct cac cag tct gat gag tct cac cat tct   336
Val Asp Asp Thr Asp Asp Ser His Gln Ser Asp Glu Ser His His Ser
            100                 105                 110

gat gaa tct gat gaa ctg gtc act gat ttt ccc acg gac ctg cca gca   384
Asp Glu Ser Asp Glu Leu Val Thr Asp Phe Pro Thr Asp Leu Pro Ala
        115                 120                 125

acc gaa gtt ttc act cca gtt gtc ccc aca gta gac aca tat gat ggc   432
Thr Glu Val Phe Thr Pro Val Val Pro Thr Val Asp Thr Tyr Asp Gly
    130                 135                 140

cga ggt gat agt gtg gtt tat gga ctg agg tca aaa tct aag aag ttt   480
Arg Gly Asp Ser Val Val Tyr Gly Leu Arg Ser Lys Ser Lys Lys Phe
```

```
145                 150                 155                 160

cgc aga cct gac atc cag tac cct gat gct aca gac gag gac atc acc   528
Arg Arg Pro Asp Ile Gln Tyr Pro Asp Ala Thr Asp Glu Asp Ile Thr
                165                 170                 175

tca cac atg gaa agc gag gag ttg aat ggt gca tac aag gcc atc ccc   576
Ser His Met Glu Ser Glu Glu Leu Asn Gly Ala Tyr Lys Ala Ile Pro
            180                 185                 190

gtt gcc cag gac ctg aac gcg cct tct gat tgg gac agc cgt ggg aag   624
Val Ala Gln Asp Leu Asn Ala Pro Ser Asp Trp Asp Ser Arg Gly Lys
        195                 200                 205

gac agt tat gaa acg agt cag ctg gat gac cag agt gct gaa acc cac   672
Asp Ser Tyr Glu Thr Ser Gln Leu Asp Asp Gln Ser Ala Glu Thr His
    210                 215                 220

agc cac aag cag tcc aga tta tat aag cgg aaa gcc aat gat gag agc   720
Ser His Lys Gln Ser Arg Leu Tyr Lys Arg Lys Ala Asn Asp Glu Ser
225                 230                 235                 240

aat gag cat tcc gat gtg att gat agt cag gaa ctt tcc aaa gtc agc   768
Asn Glu His Ser Asp Val Ile Asp Ser Gln Glu Leu Ser Lys Val Ser
                245                 250                 255

cgt gaa ttc cac agc cat gaa ttt cac agc cat gaa gat atg ctg gtt   816
Arg Glu Phe His Ser His Glu Phe His Ser His Glu Asp Met Leu Val
            260                 265                 270

gta gac ccc aaa agt aag gaa gaa gat aaa cac ctg aaa ttt cgt att   864
Val Asp Pro Lys Ser Lys Glu Glu Asp Lys His Leu Lys Phe Arg Ile
        275                 280                 285

tct cat gaa tta gat agt gca tct tct gag gtc aat taa               903
Ser His Glu Leu Asp Ser Ala Ser Ser Glu Val Asn  *
    290                 295                 300


<210> 42
<211> 300
<212> PRT
<213> human

<400> 42
Met Arg Ile Ala Val Ile Cys Phe Cys Leu Leu Gly Ile Thr Cys Ala
 1               5                  10                  15
Ile Pro Val Lys Gln Ala Asp Ser Gly Ser Ser Glu Glu Lys Gln Leu
            20                  25                  30
Tyr Asn Lys Tyr Pro Asp Ala Val Ala Thr Trp Leu Asn Pro Asp Pro
        35                  40                  45
Ser Gln Lys Gln Asn Leu Leu Ala Pro Gln Thr Leu Pro Ser Lys Ser
    50                  55                  60
Asn Glu Ser His Asp His Met Asp Asp Met Asp Asp Glu Asp Asp Asp
65                  70                  75                  80
Asp His Val Asp Ser Gln Asp Ser Ile Asp Ser Asn Asp Ser Asp Asp
                85                  90                  95
Val Asp Asp Thr Asp Asp Ser His Gln Ser Asp Glu Ser His His Ser
            100                 105                 110
Asp Glu Ser Asp Glu Leu Val Thr Asp Phe Pro Thr Asp Leu Pro Ala
        115                 120                 125
```

```
Thr Glu Val Phe Thr Pro Val Val Pro Thr Val Asp Thr Tyr Asp Gly
    130                 135                 140
Arg Gly Asp Ser Val Val Tyr Gly Leu Arg Ser Lys Ser Lys Lys Phe
145                 150                 155                 160
Arg Arg Pro Asp Ile Gln Tyr Pro Asp Ala Thr Asp Glu Asp Ile Thr
                165                 170                 175
Ser His Met Glu Ser Glu Glu Leu Asn Gly Ala Tyr Lys Ala Ile Pro
            180                 185                 190
Val Ala Gln Asp Leu Asn Ala Pro Ser Asp Trp Asp Ser Arg Gly Lys
        195                 200                 205
Asp Ser Tyr Glu Thr Ser Gln Leu Asp Asp Gln Ser Ala Glu Thr His
    210                 215                 220
Ser His Lys Gln Ser Arg Leu Tyr Lys Arg Lys Ala Asn Asp Glu Ser
225                 230                 235                 240
Asn Glu His Ser Asp Val Ile Asp Ser Gln Glu Leu Ser Lys Val Ser
                245                 250                 255
Arg Glu Phe His Ser His Glu Phe His Ser His Glu Asp Met Leu Val
            260                 265                 270
Val Asp Pro Lys Ser Lys Glu Glu Asp Lys His Leu Lys Phe Arg Ile
        275                 280                 285
Ser His Glu Leu Asp Ser Ala Ser Ser Glu Val Asn
    290                 295                 300


<210> 43
<211> 3513
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(3513)

<400> 43
atg ggg ctg gcc tgg gga cta ggc gtc ctg ttc ctg atg cat gtg tgt   48
Met Gly Leu Ala Trp Gly Leu Gly Val Leu Phe Leu Met His Val Cys
 1               5                  10                  15

ggc acc aac cgc att cca gag tct ggc gga gac aac agc gtg ttt gac   96
Gly Thr Asn Arg Ile Pro Glu Ser Gly Gly Asp Asn Ser Val Phe Asp
            20                  25                  30

atc ttt gaa ctc acc ggg gcc gcc cgc aag ggg tct ggg cgc cga ctg   144
Ile Phe Glu Leu Thr Gly Ala Ala Arg Lys Gly Ser Gly Arg Arg Leu
        35                  40                  45

gtg aag ggc ccc gac cct tcc agc cca gct ttc cgc atc gag gat gcc   192
Val Lys Gly Pro Asp Pro Ser Ser Pro Ala Phe Arg Ile Glu Asp Ala
    50                  55                  60

aac ctg atc ccc cct gtg cct gat gac aag ttc caa gac ctg gtg gat   240
Asn Leu Ile Pro Pro Val Pro Asp Asp Lys Phe Gln Asp Leu Val Asp
65                  70                  75                  80

gct gtg cgg gca gaa aag ggt ttc ctc ctt ctg gca tcc ctg agg cag   288
Ala Val Arg Ala Glu Lys Gly Phe Leu Leu Leu Ala Ser Leu Arg Gln
                85                  90                  95

atg aag aag acc cgg ggc acg ctg ctg gcc ctg gag cgg aaa gac cac   336
Met Lys Lys Thr Arg Gly Thr Leu Leu Ala Leu Glu Arg Lys Asp His
            100                 105                 110
```

```
tct ggc cag gtc ttc agc gtg gtg tcc aat ggc aag gcg ggc acc ctg   384
Ser Gly Gln Val Phe Ser Val Val Ser Asn Gly Lys Ala Gly Thr Leu
        115                 120                 125

gac ctc agc ctg acc gtc caa gga aag cag cac gtg gtg tct gtg gaa   432
Asp Leu Ser Leu Thr Val Gln Gly Lys Gln His Val Val Ser Val Glu
    130                 135                 140

gaa gct ctc ctg gca acc ggc cag tgg aag agc atc acc ctg ttt gtg   480
Glu Ala Leu Leu Ala Thr Gly Gln Trp Lys Ser Ile Thr Leu Phe Val
145                 150                 155                 160

cag gaa gac agg gcc cag ctg tac atc gac tgt gaa aag atg gag aat   528
Gln Glu Asp Arg Ala Gln Leu Tyr Ile Asp Cys Glu Lys Met Glu Asn
                165                 170                 175

gct gag ttg gac gtc ccc atc caa agc gtc ttc acc aga gac ctg gcc   576
Ala Glu Leu Asp Val Pro Ile Gln Ser Val Phe Thr Arg Asp Leu Ala
            180                 185                 190

agc atc gcc aga ctc cgc atc gca aag ggg ggc gtc aat gac aat ttc   624
Ser Ile Ala Arg Leu Arg Ile Ala Lys Gly Gly Val Asn Asp Asn Phe
        195                 200                 205

cag ggg gtg ctg cag aat gtg agg ttt gtc ttt gga acc aca cca gaa   672
Gln Gly Val Leu Gln Asn Val Arg Phe Val Phe Gly Thr Thr Pro Glu
    210                 215                 220

gac atc ctc agg aac aaa ggc tgc tcc agc tct acc agt gtc ctc ctc   720
Asp Ile Leu Arg Asn Lys Gly Cys Ser Ser Ser Thr Ser Val Leu Leu
225                 230                 235                 240

acc ctt gac aac aac gtg gtg aat ggt tcc agc cct gcc atc cgc act   768
Thr Leu Asp Asn Asn Val Val Asn Gly Ser Ser Pro Ala Ile Arg Thr
                245                 250                 255

aac tac att ggc cac aag aca aag gac ttg caa gcc atc tgc ggc atc   816
Asn Tyr Ile Gly His Lys Thr Lys Asp Leu Gln Ala Ile Cys Gly Ile
            260                 265                 270

tcc tgt gat gag ctg tcc agc atg gtc ctg gaa ctc agg ggc ctg cgc   864
Ser Cys Asp Glu Leu Ser Ser Met Val Leu Glu Leu Arg Gly Leu Arg
        275                 280                 285

acc att gtg acc acg ctg cag gac agc atc cgc aaa gtg act gaa gag   912
Thr Ile Val Thr Thr Leu Gln Asp Ser Ile Arg Lys Val Thr Glu Glu
    290                 295                 300

aac aaa gag ttg gcc aat gag ctg agg cgg cct ccc cta tgc tat cac   960
Asn Lys Glu Leu Ala Asn Glu Leu Arg Arg Pro Pro Leu Cys Tyr His
305                 310                 315                 320

aac gga gtt cag tac aga aat aac gag gaa tgg act gtt gat agc tgc  1008
Asn Gly Val Gln Tyr Arg Asn Asn Glu Glu Trp Thr Val Asp Ser Cys
                325                 330                 335

act gag tgt cac tgt cag aac tca gtt acc atc tgc aaa aag gtg tcc  1056
Thr Glu Cys His Cys Gln Asn Ser Val Thr Ile Cys Lys Lys Val Ser
            340                 345                 350
```

```
tgc ccc atc atg ccc tgc tcc aat gcc aca gtt cct gat gga gaa tgc   1104
Cys Pro Ile Met Pro Cys Ser Asn Ala Thr Val Pro Asp Gly Glu Cys
        355                 360                 365

tgt cct cgc tgt tgg ccc agc gac tct gcg gac gat ggc tgg tct cca   1152
Cys Pro Arg Cys Trp Pro Ser Asp Ser Ala Asp Asp Gly Trp Ser Pro
    370                 375                 380

tgg tcc gag tgg acc tcc tgt tct acg agc tgt ggc aat gga att cag   1200
Trp Ser Glu Trp Thr Ser Cys Ser Thr Ser Cys Gly Asn Gly Ile Gln
385                 390                 395                 400

cag cgc ggc cgc tcc tgc gat agc ctc aac aac cga tgt gag ggc tcc   1248
Gln Arg Gly Arg Ser Cys Asp Ser Leu Asn Asn Arg Cys Glu Gly Ser
                405                 410                 415

tcg gtc cag aca cgg acc tgc cac att cag gag tgt gac aaa aga ttt   1296
Ser Val Gln Thr Arg Thr Cys His Ile Gln Glu Cys Asp Lys Arg Phe
            420                 425                 430

aaa cag gat ggt ggc tgg agc cac tgg tcc ccg tgg tca tct tgt tct   1344
Lys Gln Asp Gly Gly Trp Ser His Trp Ser Pro Trp Ser Ser Cys Ser
        435                 440                 445

gtg aca tgt ggt gat ggt gtg atc aca agg atc cgg ctc tgc aac tct   1392
Val Thr Cys Gly Asp Gly Val Ile Thr Arg Ile Arg Leu Cys Asn Ser
    450                 455                 460

ccc agc ccc cag atg aat ggg aaa ccc tgt gaa ggc gaa gcg cgg gag   1440
Pro Ser Pro Gln Met Asn Gly Lys Pro Cys Glu Gly Glu Ala Arg Glu
465                 470                 475                 480

acc aaa gcc tgc aag aaa gac gcc tgc ccc atc aat gga ggc tgg ggt   1488
Thr Lys Ala Cys Lys Lys Asp Ala Cys Pro Ile Asn Gly Gly Trp Gly
                485                 490                 495

cct tgg tca cca tgg gac atc tgt tct gtc acc tgt gga gga ggg gta   1536
Pro Trp Ser Pro Trp Asp Ile Cys Ser Val Thr Cys Gly Gly Gly Val
            500                 505                 510

cag aaa cgt agt cgt ctc tgc aac aac ccc gca ccc cag ttt gga ggc   1584
Gln Lys Arg Ser Arg Leu Cys Asn Asn Pro Ala Pro Gln Phe Gly Gly
        515                 520                 525

aag gac tgc gtt ggt gat gta aca gaa aac cag atc tgc aac aag cag   1632
Lys Asp Cys Val Gly Asp Val Thr Glu Asn Gln Ile Cys Asn Lys Gln
    530                 535                 540

gac tgt cca att gat gga tgc ctg tcc aat ccc tgc ttt gcc ggc gtg   1680
Asp Cys Pro Ile Asp Gly Cys Leu Ser Asn Pro Cys Phe Ala Gly Val
545                 550                 555                 560

aag tgt act agc tac cct gat ggc agc tgg aaa tgt ggt gct tgt ccc   1728
Lys Cys Thr Ser Tyr Pro Asp Gly Ser Trp Lys Cys Gly Ala Cys Pro
                565                 570                 575

cct ggt tac agt gga aat ggc atc cag tgc aca gat gtt gat gag tgc   1776
Pro Gly Tyr Ser Gly Asn Gly Ile Gln Cys Thr Asp Val Asp Glu Cys
            580                 585                 590

aaa gaa gtg cct gat gcc tgc ttc aac cac aat gga gag cac cgg tgt   1824
```

```
Lys Glu Val Pro Asp Ala Cys Phe Asn His Asn Gly Glu His Arg Cys
        595                 600                 605

gag aac acg gac ccc ggc tac aac tgc ctg ccc tgc ccc cca cgc ttc   1872
Glu Asn Thr Asp Pro Gly Tyr Asn Cys Leu Pro Cys Pro Pro Arg Phe
    610                 615                 620

acc ggc tca cag ccc ttc ggc cag ggt gtc gaa cat gcc acg gcc aac   1920
Thr Gly Ser Gln Pro Phe Gly Gln Gly Val Glu His Ala Thr Ala Asn
625                 630                 635                 640

aaa cag gtg tgc aag ccc cgt aac ccc tgc acg gat ggg acc cac gac   1968
Lys Gln Val Cys Lys Pro Arg Asn Pro Cys Thr Asp Gly Thr His Asp
                645                 650                 655

tgc aac aag aac gcc aag tgc aac tac ctg ggc cac tat agc gac ccc   2016
Cys Asn Lys Asn Ala Lys Cys Asn Tyr Leu Gly His Tyr Ser Asp Pro
            660                 665                 670

atg tac cgc tgc gag tgc aag cct ggc tac gct ggc aat ggc atc atc   2064
Met Tyr Arg Cys Glu Cys Lys Pro Gly Tyr Ala Gly Asn Gly Ile Ile
        675                 680                 685

tgc ggg gag gac aca gac ctg gat ggc tgg ccc aat gag aac ctg gtg   2112
Cys Gly Glu Asp Thr Asp Leu Asp Gly Trp Pro Asn Glu Asn Leu Val
    690                 695                 700

tgc gtg gcc aat gcg act tac cac tgc aaa aag gat aat tgc ccc aac   2160
Cys Val Ala Asn Ala Thr Tyr His Cys Lys Lys Asp Asn Cys Pro Asn
705                 710                 715                 720

ctt ccc aac tca ggg cag gaa gac tat gac aag gat gga att ggt gat   2208
Leu Pro Asn Ser Gly Gln Glu Asp Tyr Asp Lys Asp Gly Ile Gly Asp
                725                 730                 735

gcc tgt gat gat gac gat gac aat gat aaa att cca gat gac agg gac   2256
Ala Cys Asp Asp Asp Asp Asp Asn Asp Lys Ile Pro Asp Asp Arg Asp
            740                 745                 750

aac tgt cca ttc cat tac aac cca gct cag tat gac tat gac aga gat   2304
Asn Cys Pro Phe His Tyr Asn Pro Ala Gln Tyr Asp Tyr Asp Arg Asp
        755                 760                 765

gat gtg gga gac cgc tgt gac aac tgt ccc tac aac cac aac cca gat   2352
Asp Val Gly Asp Arg Cys Asp Asn Cys Pro Tyr Asn His Asn Pro Asp
    770                 775                 780

cag gca gac aca gac aac aat ggg gaa gga gac gcc tgt gct gca gac   2400
Gln Ala Asp Thr Asp Asn Asn Gly Glu Gly Asp Ala Cys Ala Ala Asp
785                 790                 795                 800

att gat gga gac ggt atc ctc aat gaa cgg gac aac tgc cag tac gtc   2448
Ile Asp Gly Asp Gly Ile Leu Asn Glu Arg Asp Asn Cys Gln Tyr Val
                805                 810                 815

tac aat gtg gac cag aga gac act gat atg gat ggg gtt gga gat cag   2496
Tyr Asn Val Asp Gln Arg Asp Thr Asp Met Asp Gly Val Gly Asp Gln
            820                 825                 830

tgt gac aat tgc ccc ttg gaa cac aat ccg gat cag ctg gac tct gac   2544
Cys Asp Asn Cys Pro Leu Glu His Asn Pro Asp Gln Leu Asp Ser Asp
```

```
        835                 840                 845

tca gac cgc att gga gat acc tgt gac aac aat cag gat att gat gaa   2592
Ser Asp Arg Ile Gly Asp Thr Cys Asp Asn Asn Gln Asp Ile Asp Glu
    850                 855                 860

gat ggc cac cag aac aat ctg gac aac tgt ccc tat gtg ccc aat gcc   2640
Asp Gly His Gln Asn Asn Leu Asp Asn Cys Pro Tyr Val Pro Asn Ala
865                 870                 875                 880

aac cag gct gac cat gac aaa gat ggc aag gga gat gcc tgt gac cac   2688
Asn Gln Ala Asp His Asp Lys Asp Gly Lys Gly Asp Ala Cys Asp His
                885                 890                 895

gat gat gac aac gat ggc att cct gat gac aag gac aac tgc aga ctc   2736
Asp Asp Asp Asn Asp Gly Ile Pro Asp Asp Lys Asp Asn Cys Arg Leu
            900                 905                 910

gtg ccc aat ccc gac cag aag gac tct gac ggc gat ggt cga ggt gat   2784
Val Pro Asn Pro Asp Gln Lys Asp Ser Asp Gly Asp Gly Arg Gly Asp
        915                 920                 925

gcc tgc aaa gat gat ttt gac cat gac agt gtg cca gac atc gat gac   2832
Ala Cys Lys Asp Asp Phe Asp His Asp Ser Val Pro Asp Ile Asp Asp
    930                 935                 940

atc tgt cct gag aat gtt gac atc agt gag acc gat ttc cgc cga ttc   2880
Ile Cys Pro Glu Asn Val Asp Ile Ser Glu Thr Asp Phe Arg Arg Phe
945                 950                 955                 960

cag atg att cct ctg gac ccc aaa ggg aca tcc caa aat gac cct aac   2928
Gln Met Ile Pro Leu Asp Pro Lys Gly Thr Ser Gln Asn Asp Pro Asn
                965                 970                 975

tgg gtt gta cgc cat cag ggt aaa gaa ctc gtc cag act gtc aac tgt   2976
Trp Val Val Arg His Gln Gly Lys Glu Leu Val Gln Thr Val Asn Cys
            980                 985                 990

gat cct gga ctc gct gta ggt tat gat gag ttt aat gct gtg gac ttc   3024
Asp Pro Gly Leu Ala Val Gly Tyr Asp Glu Phe Asn Ala Val Asp Phe
        995                 1000                1005

agt ggc acc ttc ttc atc aac acc gaa agg gac gat gac tat gct gga   3072
Ser Gly Thr Phe Phe Ile Asn Thr Glu Arg Asp Asp Asp Tyr Ala Gly
    1010                1015                1020

ttt gtc ttt ggc tac cag tcc agc agc cgc ttt tat gtt gtg atg tgg   3120
Phe Val Phe Gly Tyr Gln Ser Ser Ser Arg Phe Tyr Val Val Met Trp
1025                1030                1035                1040

aag caa gtc acc cag tcc tac tgg gac acc aac ccc acg agg gct cag   3168
Lys Gln Val Thr Gln Ser Tyr Trp Asp Thr Asn Pro Thr Arg Ala Gln
                1045                1050                1055

gga tac tcg ggc ctt tct gtg aaa gtt gta aac tcc acc aca ggg cct   3216
Gly Tyr Ser Gly Leu Ser Val Lys Val Val Asn Ser Thr Thr Gly Pro
            1060                1065                1070

ggc gag cac ctg cgg aac gcc ctg tgg cac aca gga aac acc cct ggc   3264
Gly Glu His Leu Arg Asn Ala Leu Trp His Thr Gly Asn Thr Pro Gly
        1075                1080                1085
```

```
cag gtg cgc acc ctg tgg cat gac cct cgt cac ata ggc tgg aaa gat   3312
Gln Val Arg Thr Leu Trp His Asp Pro Arg His Ile Gly Trp Lys Asp
    1090                1095                1100

ttc acc gcc tac aga tgg cgt ctc agc cac agg cca aag acg ggt ttc   3360
Phe Thr Ala Tyr Arg Trp Arg Leu Ser His Arg Pro Lys Thr Gly Phe
1105                1110                1115                1120

att aga gtg gtg atg tat gaa ggg aag aaa atc atg gct gac tca gga   3408
Ile Arg Val Val Met Tyr Glu Gly Lys Lys Ile Met Ala Asp Ser Gly
                1125                1130                1135

ccc atc tat gat aaa acc tat gct ggt ggt aga cta ggg ttg ttt gtc   3456
Pro Ile Tyr Asp Lys Thr Tyr Ala Gly Gly Arg Leu Gly Leu Phe Val
            1140                1145                1150

ttc tct caa gaa atg gtg ttc ttc tct gac ctg aaa tac gaa tgt aga   3504
Phe Ser Gln Glu Met Val Phe Phe Ser Asp Leu Lys Tyr Glu Cys Arg
        1155                1160                1165

gat ccc taa                                                      3513
Asp Pro  *
    1170
```

<210> 44
<211> 1170
<212> PRT
<213> human

<400> 44

```
Met Gly Leu Ala Trp Gly Leu Gly Val Leu Phe Leu Met His Val Cys
 1               5                  10                  15
Gly Thr Asn Arg Ile Pro Glu Ser Gly Gly Asp Asn Ser Val Phe Asp
            20                  25                  30
Ile Phe Glu Leu Thr Gly Ala Ala Arg Lys Gly Ser Gly Arg Arg Leu
        35                  40                  45
Val Lys Gly Pro Asp Pro Ser Ser Pro Ala Phe Arg Ile Glu Asp Ala
    50                  55                  60
Asn Leu Ile Pro Pro Val Pro Asp Asp Lys Phe Gln Asp Leu Val Asp
65                  70                  75                  80
Ala Val Arg Ala Glu Lys Gly Phe Leu Leu Leu Ala Ser Leu Arg Gln
                85                  90                  95
Met Lys Lys Thr Arg Gly Thr Leu Leu Ala Leu Glu Arg Lys Asp His
            100                 105                 110
Ser Gly Gln Val Phe Ser Val Val Ser Asn Gly Lys Ala Gly Thr Leu
        115                 120                 125
Asp Leu Ser Leu Thr Val Gln Gly Lys Gln His Val Val Ser Val Glu
    130                 135                 140
Glu Ala Leu Leu Ala Thr Gly Gln Trp Lys Ser Ile Thr Leu Phe Val
145                 150                 155                 160
Gln Glu Asp Arg Ala Gln Leu Tyr Ile Asp Cys Glu Lys Met Glu Asn
                165                 170                 175
Ala Glu Leu Asp Val Pro Ile Gln Ser Val Phe Thr Arg Asp Leu Ala
            180                 185                 190
Ser Ile Ala Arg Leu Arg Ile Ala Lys Gly Gly Val Asn Asp Asn Phe
        195                 200                 205
Gln Gly Val Leu Gln Asn Val Arg Phe Val Phe Gly Thr Thr Pro Glu
    210                 215                 220
Asp Ile Leu Arg Asn Lys Gly Cys Ser Ser Ser Thr Ser Val Leu Leu
```

```
225                 230                 235                 240
Thr Leu Asp Asn Asn Val Val Asn Gly Ser Ser Pro Ala Ile Arg Thr
                245                 250                 255
Asn Tyr Ile Gly His Lys Thr Lys Asp Leu Gln Ala Ile Cys Gly Ile
            260                 265                 270
Ser Cys Asp Glu Leu Ser Ser Met Val Leu Glu Leu Arg Gly Leu Arg
        275                 280                 285
Thr Ile Val Thr Thr Leu Gln Asp Ser Ile Arg Lys Val Thr Glu Glu
    290                 295                 300
Asn Lys Glu Leu Ala Asn Glu Leu Arg Arg Pro Pro Leu Cys Tyr His
305                 310                 315                 320
Asn Gly Val Gln Tyr Arg Asn Asn Glu Glu Trp Thr Val Asp Ser Cys
                325                 330                 335
Thr Glu Cys His Cys Gln Asn Ser Val Thr Ile Cys Lys Lys Val Ser
            340                 345                 350
Cys Pro Ile Met Pro Cys Ser Asn Ala Thr Val Pro Asp Gly Glu Cys
        355                 360                 365
Cys Pro Arg Cys Trp Pro Ser Asp Ser Ala Asp Asp Gly Trp Ser Pro
    370                 375                 380
Trp Ser Glu Trp Thr Ser Cys Ser Thr Ser Cys Gly Asn Gly Ile Gln
385                 390                 395                 400
Gln Arg Gly Arg Ser Cys Asp Ser Leu Asn Asn Arg Cys Glu Gly Ser
                405                 410                 415
Ser Val Gln Thr Arg Thr Cys His Ile Gln Glu Cys Asp Lys Arg Phe
            420                 425                 430
Lys Gln Asp Gly Gly Trp Ser His Trp Ser Pro Trp Ser Ser Cys Ser
        435                 440                 445
Val Thr Cys Gly Asp Gly Val Ile Thr Arg Ile Arg Leu Cys Asn Ser
    450                 455                 460
Pro Ser Pro Gln Met Asn Gly Lys Pro Cys Glu Gly Glu Ala Arg Glu
465                 470                 475                 480
Thr Lys Ala Cys Lys Lys Asp Ala Cys Pro Ile Asn Gly Gly Trp Gly
                485                 490                 495
Pro Trp Ser Pro Trp Asp Ile Cys Ser Val Thr Cys Gly Gly Gly Val
            500                 505                 510
Gln Lys Arg Ser Arg Leu Cys Asn Asn Pro Ala Pro Gln Phe Gly Gly
        515                 520                 525
Lys Asp Cys Val Gly Asp Val Thr Glu Asn Gln Ile Cys Asn Lys Gln
    530                 535                 540
Asp Cys Pro Ile Asp Gly Cys Leu Ser Asn Pro Cys Phe Ala Gly Val
545                 550                 555                 560
Lys Cys Thr Ser Tyr Pro Asp Gly Ser Trp Lys Cys Gly Ala Cys Pro
                565                 570                 575
Pro Gly Tyr Ser Gly Asn Gly Ile Gln Cys Thr Asp Val Asp Glu Cys
            580                 585                 590
Lys Glu Val Pro Asp Ala Cys Phe Asn His Asn Gly Glu His Arg Cys
        595                 600                 605
Glu Asn Thr Asp Pro Gly Tyr Asn Cys Leu Pro Cys Pro Pro Arg Phe
    610                 615                 620
Thr Gly Ser Gln Pro Phe Gly Gln Gly Val Glu His Ala Thr Ala Asn
625                 630                 635                 640
Lys Gln Val Cys Lys Pro Arg Asn Pro Cys Thr Asp Gly Thr His Asp
                645                 650                 655
Cys Asn Lys Asn Ala Lys Cys Asn Tyr Leu Gly His Tyr Ser Asp Pro
            660                 665                 670
Met Tyr Arg Cys Glu Cys Lys Pro Gly Tyr Ala Gly Asn Gly Ile Ile
        675                 680                 685
Cys Gly Glu Asp Thr Asp Leu Asp Gly Trp Pro Asn Glu Asn Leu Val
    690                 695                 700
Cys Val Ala Asn Ala Thr Tyr His Cys Lys Lys Asp Asn Cys Pro Asn
705                 710                 715                 720
```

```
Leu Pro Asn Ser Gly Gln Glu Asp Tyr Asp Lys Asp Gly Ile Gly Asp
                725                 730                 735
Ala Cys Asp Asp Asp Asp Asp Asn Asp Lys Ile Pro Asp Asp Arg Asp
            740                 745                 750
Asn Cys Pro Phe His Tyr Asn Pro Ala Gln Tyr Asp Tyr Asp Arg Asp
        755                 760                 765
Asp Val Gly Asp Arg Cys Asp Asn Cys Pro Tyr Asn His Asn Pro Asp
    770                 775                 780
Gln Ala Asp Thr Asp Asn Asn Gly Glu Gly Asp Ala Cys Ala Ala Asp
785                 790                 795                 800
Ile Asp Gly Asp Gly Ile Leu Asn Glu Arg Asp Asn Cys Gln Tyr Val
                805                 810                 815
Tyr Asn Val Asp Gln Arg Asp Thr Asp Met Asp Gly Val Gly Asp Gln
            820                 825                 830
Cys Asp Asn Cys Pro Leu Glu His Asn Pro Asp Gln Leu Asp Ser Asp
        835                 840                 845
Ser Asp Arg Ile Gly Asp Thr Cys Asp Asn Asn Gln Asp Ile Asp Glu
    850                 855                 860
Asp Gly His Gln Asn Asn Leu Asp Asn Cys Pro Tyr Val Pro Asn Ala
865                 870                 875                 880
Asn Gln Ala Asp His Asp Lys Asp Gly Lys Gly Asp Ala Cys Asp His
                885                 890                 895
Asp Asp Asp Asn Asp Gly Ile Pro Asp Asp Lys Asp Asn Cys Arg Leu
            900                 905                 910
Val Pro Asn Pro Asp Gln Lys Asp Ser Asp Gly Asp Gly Arg Gly Asp
        915                 920                 925
Ala Cys Lys Asp Asp Phe Asp His Asp Ser Val Pro Asp Ile Asp Asp
    930                 935                 940
Ile Cys Pro Glu Asn Val Asp Ile Ser Glu Thr Asp Phe Arg Arg Phe
945                 950                 955                 960
Gln Met Ile Pro Leu Asp Pro Lys Gly Thr Ser Gln Asn Asp Pro Asn
                965                 970                 975
Trp Val Val Arg His Gln Gly Lys Glu Leu Val Gln Thr Val Asn Cys
            980                 985                 990
Asp Pro Gly Leu Ala Val Gly Tyr Asp Glu Phe Asn Ala Val Asp Phe
        995                 1000                1005
Ser Gly Thr Phe Phe Ile Asn Thr Glu Arg Asp Asp Asp Tyr Ala Gly
    1010                1015                1020
Phe Val Phe Gly Tyr Gln Ser Ser Ser Arg Phe Tyr Val Val Met Trp
1025                1030                1035                1040
Lys Gln Val Thr Gln Ser Tyr Trp Asp Thr Asn Pro Thr Arg Ala Gln
                1045                1050                1055
Gly Tyr Ser Gly Leu Ser Val Lys Val Val Asn Ser Thr Thr Gly Pro
            1060                1065                1070
Gly Glu His Leu Arg Asn Ala Leu Trp His Thr Gly Asn Thr Pro Gly
        1075                1080                1085
Gln Val Arg Thr Leu Trp His Asp Pro Arg His Ile Gly Trp Lys Asp
    1090                1095                1100
Phe Thr Ala Tyr Arg Trp Arg Leu Ser His Arg Pro Lys Thr Gly Phe
1105                1110                1115                1120
Ile Arg Val Val Met Tyr Glu Gly Lys Lys Ile Met Ala Asp Ser Gly
                1125                1130                1135
Pro Ile Tyr Asp Lys Thr Tyr Ala Gly Gly Arg Leu Gly Leu Phe Val
            1140                1145                1150
Phe Ser Gln Glu Met Val Phe Phe Ser Asp Leu Lys Tyr Glu Cys Arg
        1155                1160                1165
Asp Pro
    1170
```

<210> 45

<211> 948
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(948)

<400> 45

```
atg gac ctg cga cag ttt ctt atg tgc ctg tcc ctg tgc aca gcc ttt   48
Met Asp Leu Arg Gln Phe Leu Met Cys Leu Ser Leu Cys Thr Ala Phe
 1               5                  10                  15

gcc ttg agc aaa ccc aca gaa aag aag gac cgt gta cat cat gag cct   96
Ala Leu Ser Lys Pro Thr Glu Lys Lys Asp Arg Val His His Glu Pro
            20                  25                  30

cag ctc agt gac aag gtt cac aat gat gct cag agt ttt gat tat gac   144
Gln Leu Ser Asp Lys Val His Asn Asp Ala Gln Ser Phe Asp Tyr Asp
        35                  40                  45

cat gat gcc ttc ttg ggt gct gaa gaa gca aag acc ttt gat cag ctg   192
His Asp Ala Phe Leu Gly Ala Glu Glu Ala Lys Thr Phe Asp Gln Leu
    50                  55                  60

aca cca gaa gag agc aag gaa agg ctt gga atg att gta gat aaa ata   240
Thr Pro Glu Glu Ser Lys Glu Arg Leu Gly Met Ile Val Asp Lys Ile
65                  70                  75                  80

gac gcg gat aaa gat ggg ttt gtg acg gag ggg gag ctg aaa tcc tgg   288
Asp Ala Asp Lys Asp Gly Phe Val Thr Glu Gly Glu Leu Lys Ser Trp
                85                  90                  95

att aag cac gcc cag aag aaa tac ata tat gac aat gtt gaa aac caa   336
Ile Lys His Ala Gln Lys Lys Tyr Ile Tyr Asp Asn Val Glu Asn Gln
            100                 105                 110

tgg cag gag ttt gat atg aat caa gac ggc tta atc tcc tgg gat gag   384
Trp Gln Glu Phe Asp Met Asn Gln Asp Gly Leu Ile Ser Trp Asp Glu
        115                 120                 125

tac aga aac gtg act tat ggc act tac ctg gat gat cca gat cct gat   432
Tyr Arg Asn Val Thr Tyr Gly Thr Tyr Leu Asp Asp Pro Asp Pro Asp
    130                 135                 140

gat gga ttt aac tat aaa cag atg atg gtt aga gat gag cgg agg ttt   480
Asp Gly Phe Asn Tyr Lys Gln Met Met Val Arg Asp Glu Arg Arg Phe
145                 150                 155                 160

aaa atg gca gac aag gat gga gac ctc att gcc acc aag gag gag ttc   528
Lys Met Ala Asp Lys Asp Gly Asp Leu Ile Ala Thr Lys Glu Glu Phe
                165                 170                 175

aca gct ttc ctg cac cct gag gag tat gac tac atg aaa gat ata gta   576
Thr Ala Phe Leu His Pro Glu Glu Tyr Asp Tyr Met Lys Asp Ile Val
            180                 185                 190

gta cag gaa aca atg gaa gat ata gat aag aat gct gat ggt ttc att   624
Val Gln Glu Thr Met Glu Asp Ile Asp Lys Asn Ala Asp Gly Phe Ile
        195                 200                 205
```

```
gat cta gaa gag tat att ggt gac atg tac agc cat gat ggg aat act   672
Asp Leu Glu Glu Tyr Ile Gly Asp Met Tyr Ser His Asp Gly Asn Thr
    210                 215                 220

gat gag cca gaa tgg gta aag aca gag cga gag cag ttt gtt gag ttt   720
Asp Glu Pro Glu Trp Val Lys Thr Glu Arg Glu Gln Phe Val Glu Phe
225                 230                 235                 240

cgg gat aag aac cgt gat ggg aag atg gac aag gaa gag acc aaa gac   768
Arg Asp Lys Asn Arg Asp Gly Lys Met Asp Lys Glu Glu Thr Lys Asp
                245                 250                 255

tgg atc ctt ccc tca gac tat gat cat gca gag gca gaa gcc agg cac   816
Trp Ile Leu Pro Ser Asp Tyr Asp His Ala Glu Ala Glu Ala Arg His
            260                 265                 270

ctg gtc tat gaa tca gac caa aac aag gat ggc aag ctt acc aag gag   864
Leu Val Tyr Glu Ser Asp Gln Asn Lys Asp Gly Lys Leu Thr Lys Glu
        275                 280                 285

gag atc gtt gac aag tat gac tta ttt gtt ggc agc cag gcc aca gat   912
Glu Ile Val Asp Lys Tyr Asp Leu Phe Val Gly Ser Gln Ala Thr Asp
    290                 295                 300

ttt ggg gag gcc tta gta cgg cat gat gag ttc tga                   948
Phe Gly Glu Ala Leu Val Arg His Asp Glu Phe  *
305                 310                 315
```

<210> 46
<211> 315
<212> PRT
<213> human

<400> 46

```
Met Asp Leu Arg Gln Phe Leu Met Cys Leu Ser Leu Cys Thr Ala Phe
 1               5                  10                  15
Ala Leu Ser Lys Pro Thr Glu Lys Lys Asp Arg Val His His Glu Pro
            20                  25                  30
Gln Leu Ser Asp Lys Val His Asn Asp Ala Gln Ser Phe Asp Tyr Asp
        35                  40                  45
His Asp Ala Phe Leu Gly Ala Glu Glu Ala Lys Thr Phe Asp Gln Leu
    50                  55                  60
Thr Pro Glu Glu Ser Lys Glu Arg Leu Gly Met Ile Val Asp Lys Ile
65                  70                  75                  80
Asp Ala Asp Lys Asp Gly Phe Val Thr Glu Gly Glu Leu Lys Ser Trp
                85                  90                  95
Ile Lys His Ala Gln Lys Lys Tyr Ile Tyr Asp Asn Val Glu Asn Gln
            100                 105                 110
Trp Gln Glu Phe Asp Met Asn Gln Asp Gly Leu Ile Ser Trp Asp Glu
        115                 120                 125
Tyr Arg Asn Val Thr Tyr Gly Thr Tyr Leu Asp Asp Pro Asp Pro Asp
    130                 135                 140
Asp Gly Phe Asn Tyr Lys Gln Met Met Val Arg Asp Glu Arg Arg Phe
145                 150                 155                 160
Lys Met Ala Asp Lys Asp Gly Asp Leu Ile Ala Thr Lys Glu Glu Phe
                165                 170                 175
Thr Ala Phe Leu His Pro Glu Glu Tyr Asp Tyr Met Lys Asp Ile Val
            180                 185                 190
Val Gln Glu Thr Met Glu Asp Ile Asp Lys Asn Ala Asp Gly Phe Ile
        195                 200                 205
```

```
Asp Leu Glu Glu Tyr Ile Gly Asp Met Tyr Ser His Asp Gly Asn Thr
    210                 215                 220
Asp Glu Pro Glu Trp Val Lys Thr Glu Arg Glu Gln Phe Val Glu Phe
225                 230                 235                 240
Arg Asp Lys Asn Arg Asp Gly Lys Met Asp Lys Glu Glu Thr Lys Asp
                245                 250                 255
Trp Ile Leu Pro Ser Asp Tyr Asp His Ala Glu Ala Glu Ala Arg His
            260                 265                 270
Leu Val Tyr Glu Ser Asp Gln Asn Lys Asp Gly Lys Leu Thr Lys Glu
        275                 280                 285
Glu Ile Val Asp Lys Tyr Asp Leu Phe Val Gly Ser Gln Ala Thr Asp
    290                 295                 300
Phe Gly Glu Ala Leu Val Arg His Asp Glu Phe
305                 310                 315


<210> 47
<211> 783
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(783)

<400> 47
atg tcc cat cac tgg ggg tac ggc aaa cac aac gga cct gag cac tgg   48
Met Ser His His Trp Gly Tyr Gly Lys His Asn Gly Pro Glu His Trp
 1               5                  10                  15

cat aag gac ttc ccc att gcc aag gga gag cgc cag tcc cct gtt gac   96
His Lys Asp Phe Pro Ile Ala Lys Gly Glu Arg Gln Ser Pro Val Asp
            20                  25                  30

atc gac act cat aca gcc aag tat gac cct tcc ctg aag ccc ctg tct   144
Ile Asp Thr His Thr Ala Lys Tyr Asp Pro Ser Leu Lys Pro Leu Ser
        35                  40                  45

gtt tcc tat gat caa gca act tcc ctg agg atc ctc aac aat ggt cat   192
Val Ser Tyr Asp Gln Ala Thr Ser Leu Arg Ile Leu Asn Asn Gly His
    50                  55                  60

gct ttc aac gtg gag ttt gat gac tct cag gac aaa gca gtg ctc aag   240
Ala Phe Asn Val Glu Phe Asp Asp Ser Gln Asp Lys Ala Val Leu Lys
65                  70                  75                  80

gga gga ccc ctg gat ggc act tac aga ttg att cag ttt cac ttt cac   288
Gly Gly Pro Leu Asp Gly Thr Tyr Arg Leu Ile Gln Phe His Phe His
                85                  90                  95

tgg ggt tca ctt gat gga caa ggt tca gag cat act gtg gat aaa aag   336
Trp Gly Ser Leu Asp Gly Gln Gly Ser Glu His Thr Val Asp Lys Lys
            100                 105                 110

aaa tat gct gca gaa ctt cac ttg gtt cac tgg aac acc aaa tat ggg   384
Lys Tyr Ala Ala Glu Leu His Leu Val His Trp Asn Thr Lys Tyr Gly
        115                 120                 125

gat ttt ggg aaa gct gtg cag caa cct gat gga ctg gcc gtt cta ggt   432
Asp Phe Gly Lys Ala Val Gln Gln Pro Asp Gly Leu Ala Val Leu Gly
    130                 135                 140
```

```
att ttt ttg aag gtt ggc agc gct aaa ccg ggc ctt cag aaa gtt gtt   480
Ile Phe Leu Lys Val Gly Ser Ala Lys Pro Gly Leu Gln Lys Val Val
145                 150                 155                 160

gat gtg ctg gat tcc att aaa aca aag ggc aag agt gct gac ttc aca   528
Asp Val Leu Asp Ser Ile Lys Thr Lys Gly Lys Ser Ala Asp Phe Thr
                165                 170                 175

aac ttt gca gct cgt ggc ctc ctt cct gaa tcc ctg gat tac tgg acc   576
Asn Phe Ala Ala Arg Gly Leu Leu Pro Glu Ser Leu Asp Tyr Trp Thr
            180                 185                 190

tac cca ggc tca ctg acc acc cct cct ctt ctg gaa tgt gtg acc tgg   624
Tyr Pro Gly Ser Leu Thr Thr Pro Pro Leu Leu Glu Cys Val Thr Trp
        195                 200                 205

att gtg ctc aag gaa ccc atc agc gtc agc agc gag cag gtg ttg aaa   672
Ile Val Leu Lys Glu Pro Ile Ser Val Ser Ser Glu Gln Val Leu Lys
    210                 215                 220

ttc cgt aaa ctt aac ttc aat ggg gag ggt gaa ccc gaa gaa ctg atg   720
Phe Arg Lys Leu Asn Phe Asn Gly Glu Gly Glu Pro Glu Glu Leu Met
225                 230                 235                 240

gtg gac aac tgg cgc cca gct cag cca ctg aag aac agg caa atc aaa   768
Val Asp Asn Trp Arg Pro Ala Gln Pro Leu Lys Asn Arg Gln Ile Lys
                245                 250                 255

gct tcc ttc aaa taa                                             783
Ala Ser Phe Lys  *
            260
```

<210> 48
<211> 260
<212> PRT
<213> human

<400> 48

```
Met Ser His His Trp Gly Tyr Gly Lys His Asn Gly Pro Glu His Trp
 1               5                  10                  15
His Lys Asp Phe Pro Ile Ala Lys Gly Glu Arg Gln Ser Pro Val Asp
            20                  25                  30
Ile Asp Thr His Thr Ala Lys Tyr Asp Pro Ser Leu Lys Pro Leu Ser
        35                  40                  45
Val Ser Tyr Asp Gln Ala Thr Ser Leu Arg Ile Leu Asn Asn Gly His
    50                  55                  60
Ala Phe Asn Val Glu Phe Asp Asp Ser Gln Asp Lys Ala Val Leu Lys
65                  70                  75                  80
Gly Gly Pro Leu Asp Gly Thr Tyr Arg Leu Ile Gln Phe His Phe His
                85                  90                  95
Trp Gly Ser Leu Asp Gly Gln Gly Ser Glu His Thr Val Asp Lys Lys
            100                 105                 110
Lys Tyr Ala Ala Glu Leu His Leu Val His Trp Asn Thr Lys Tyr Gly
        115                 120                 125
Asp Phe Gly Lys Ala Val Gln Gln Pro Asp Gly Leu Ala Val Leu Gly
    130                 135                 140
Ile Phe Leu Lys Val Gly Ser Ala Lys Pro Gly Leu Gln Lys Val Val
145                 150                 155                 160
Asp Val Leu Asp Ser Ile Lys Thr Lys Gly Lys Ser Ala Asp Phe Thr
```

```
                165                 170                 175
Asn Phe Ala Ala Arg Gly Leu Leu Pro Glu Ser Leu Asp Tyr Trp Thr
            180                 185                 190
Tyr Pro Gly Ser Leu Thr Thr Pro Pro Leu Leu Glu Cys Val Thr Trp
        195                 200                 205
Ile Val Leu Lys Glu Pro Ile Ser Val Ser Ser Glu Gln Val Leu Lys
    210                 215                 220
Phe Arg Lys Leu Asn Phe Asn Gly Glu Gly Glu Pro Glu Glu Leu Met
225                 230                 235                 240
Val Asp Asn Trp Arg Pro Ala Gln Pro Leu Lys Asn Arg Gln Ile Lys
                245                 250                 255
Ala Ser Phe Lys
            260
```

<210> 49
<211> 897
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(897)

<400> 49

```
atg aag act ctg cag tct aca ctt ctc ctg tta ctg ctt gtg cct ctg   48
Met Lys Thr Leu Gln Ser Thr Leu Leu Leu Leu Leu Val Pro Leu
 1               5                  10                  15

ata aag cca gca cca cca acc cag cag gac tca cgc att atc tat gat   96
Ile Lys Pro Ala Pro Pro Thr Gln Gln Asp Ser Arg Ile Ile Tyr Asp
            20                  25                  30

tat gga aca gat aat ttt gaa gaa tcc ata ttt agc caa gat tat gag   144
Tyr Gly Thr Asp Asn Phe Glu Glu Ser Ile Phe Ser Gln Asp Tyr Glu
        35                  40                  45

gat aaa tac ctg gat gga aaa aat att aag gaa aaa gaa act gtg ata   192
Asp Lys Tyr Leu Asp Gly Lys Asn Ile Lys Glu Lys Glu Thr Val Ile
    50                  55                  60

ata ccc aat gag aaa agt ctt caa tta caa aaa gat gag gca ata aca   240
Ile Pro Asn Glu Lys Ser Leu Gln Leu Gln Lys Asp Glu Ala Ile Thr
65                  70                  75                  80

cca tta cct ccc aag aaa gaa aat gat gaa atg ccc acg tgt ctg ctg   288
Pro Leu Pro Pro Lys Lys Glu Asn Asp Glu Met Pro Thr Cys Leu Leu
                85                  90                  95

tgt gtt tgt tta agt ggc tct gta tac tgt gaa gaa gtt gac att gat   336
Cys Val Cys Leu Ser Gly Ser Val Tyr Cys Glu Glu Val Asp Ile Asp
            100                 105                 110

gct gta cca ccc tta cca aag gaa tca gcc tat ctt tac gca cga ttc   384
Ala Val Pro Pro Leu Pro Lys Glu Ser Ala Tyr Leu Tyr Ala Arg Phe
        115                 120                 125

aac aaa att aaa aag ctg act gcc aaa gat ttt gca gac ata cct aac   432
Asn Lys Ile Lys Lys Leu Thr Ala Lys Asp Phe Ala Asp Ile Pro Asn
    130                 135                 140
```

```
tta aga aga ctc gat ttt aca gga aat ttg ata gaa gat ata gaa gat   480
Leu Arg Arg Leu Asp Phe Thr Gly Asn Leu Ile Glu Asp Ile Glu Asp
145                 150                 155                 160

ggt act ttt tca aaa ctt tct ctg tta gaa gaa ctt tca ctt gct gaa   528
Gly Thr Phe Ser Lys Leu Ser Leu Leu Glu Glu Leu Ser Leu Ala Glu
                165                 170                 175

aat caa cta cta aaa ctt cca gtt ctt cct ccc aag ctc act tta ttt   576
Asn Gln Leu Leu Lys Leu Pro Val Leu Pro Pro Lys Leu Thr Leu Phe
            180                 185                 190

aat gca aaa tac aac aaa atc aag agt agg gga atc aaa gca aat gca   624
Asn Ala Lys Tyr Asn Lys Ile Lys Ser Arg Gly Ile Lys Ala Asn Ala
        195                 200                 205

ttc aaa aaa ctg aat aac ctc acc ttc ctc tac ttg gac cat aat gcc   672
Phe Lys Lys Leu Asn Asn Leu Thr Phe Leu Tyr Leu Asp His Asn Ala
    210                 215                 220

ctg gaa tcc gtg cct ctt aat tta cca gaa agt cta cgt gta att cat   720
Leu Glu Ser Val Pro Leu Asn Leu Pro Glu Ser Leu Arg Val Ile His
225                 230                 235                 240

ctt cag ttc aac aac ata gct tca att aca gat gac aca ttc tgc aag   768
Leu Gln Phe Asn Asn Ile Ala Ser Ile Thr Asp Asp Thr Phe Cys Lys
                245                 250                 255

gct aat gac acc agt tac atc cgg gac cgc att gaa gag ata cgc ctg   816
Ala Asn Asp Thr Ser Tyr Ile Arg Asp Arg Ile Glu Glu Ile Arg Leu
            260                 265                 270

gag ggc aat cca atc gtc ctg gga aag cat cca aac agt ttt att tgc   864
Glu Gly Asn Pro Ile Val Leu Gly Lys His Pro Asn Ser Phe Ile Cys
        275                 280                 285

tta aaa aga tta ccg ata ggg tca tac ttt taa                      897
Leu Lys Arg Leu Pro Ile Gly Ser Tyr Phe  *
    290                 295
```

<210> 50
<211> 298
<212> PRT
<213> human

<400> 50

```
Met Lys Thr Leu Gln Ser Thr Leu Leu Leu Leu Leu Leu Val Pro Leu
 1               5                  10                  15
Ile Lys Pro Ala Pro Pro Thr Gln Gln Asp Ser Arg Ile Ile Tyr Asp
            20                  25                  30
Tyr Gly Thr Asp Asn Phe Glu Glu Ser Ile Phe Ser Gln Asp Tyr Glu
        35                  40                  45
Asp Lys Tyr Leu Asp Gly Lys Asn Ile Lys Glu Lys Glu Thr Val Ile
    50                  55                  60
Ile Pro Asn Glu Lys Ser Leu Gln Leu Gln Lys Asp Glu Ala Ile Thr
65                  70                  75                  80
Pro Leu Pro Pro Lys Lys Glu Asn Asp Glu Met Pro Thr Cys Leu Leu
                85                  90                  95
Cys Val Cys Leu Ser Gly Ser Val Tyr Cys Glu Glu Val Asp Ile Asp
            100                 105                 110
```

```
Ala Val Pro Pro Leu Pro Lys Glu Ser Ala Tyr Leu Tyr Ala Arg Phe
        115                 120                 125
Asn Lys Ile Lys Lys Leu Thr Ala Lys Asp Phe Ala Asp Ile Pro Asn
    130                 135                 140
Leu Arg Arg Leu Asp Phe Thr Gly Asn Leu Ile Glu Asp Ile Glu Asp
145                 150                 155                 160
Gly Thr Phe Ser Lys Leu Ser Leu Leu Glu Glu Leu Ser Leu Ala Glu
                165                 170                 175
Asn Gln Leu Leu Lys Leu Pro Val Leu Pro Pro Lys Leu Thr Leu Phe
            180                 185                 190
Asn Ala Lys Tyr Asn Lys Ile Lys Ser Arg Gly Ile Lys Ala Asn Ala
        195                 200                 205
Phe Lys Lys Leu Asn Asn Leu Thr Phe Leu Tyr Leu Asp His Asn Ala
    210                 215                 220
Leu Glu Ser Val Pro Leu Asn Leu Pro Glu Ser Leu Arg Val Ile His
225                 230                 235                 240
Leu Gln Phe Asn Asn Ile Ala Ser Ile Thr Asp Asp Thr Phe Cys Lys
                245                 250                 255
Ala Asn Asp Thr Ser Tyr Ile Arg Asp Arg Ile Glu Glu Ile Arg Leu
            260                 265                 270
Glu Gly Asn Pro Ile Val Leu Gly Lys His Pro Asn Ser Phe Ile Cys
        275                 280                 285
Leu Lys Arg Leu Pro Ile Gly Ser Tyr Phe
    290                 295
```

<210> 51
<211> 369
<212> DNA
<213> human

<220>
<221> CDS
<222> (1)...(369)

<400> 51

```
atg gaa gtg gaa cca gct gcg atg gtg aaa gtg act ctc tcc ccc cca   48
Met Glu Val Glu Pro Ala Ala Met Val Lys Val Thr Leu Ser Pro Pro
 1               5                  10                  15

ctc ttt tct ctg cca gaa ttt tgg aac ctg gct ggg cca cct gct ccg   96
Leu Phe Ser Leu Pro Glu Phe Trp Asn Leu Ala Gly Pro Pro Ala Pro
             20                  25                  30

gta cac act gcc aca ggg agg gtc acc gtc att ggt atc tgg cag gat   144
Val His Thr Ala Thr Gly Arg Val Thr Val Ile Gly Ile Trp Gln Asp
         35                  40                  45

aga cga gat tac atc tgt ccc tac tgt cag ctg aag gcc gag ctc tgc   192
Arg Arg Asp Tyr Ile Cys Pro Tyr Cys Gln Leu Lys Ala Glu Leu Cys
     50                  55                  60

cag caa atc agc agg gaa gct gtt gag cag cct ctg gca ggt caa aaa   240
Gln Gln Ile Ser Arg Glu Ala Val Glu Gln Pro Leu Ala Gly Gln Lys
 65                  70                  75                  80

gct gcc ttc aaa aca gct caa acg gct cct gca aca ggc cca atg aca   288
Ala Ala Phe Lys Thr Ala Gln Thr Ala Pro Ala Thr Gly Pro Met Thr
                 85                  90                  95

gca gca gcg tgt gag cac aag gca gtg gga gct gcg gca gcc ttt ttg   336
```

```
Ala Ala Ala Cys Glu His Lys Ala Val Gly Ala Ala Ala Ala Phe Leu
            100                 105                 110

cct caa gga gga gac agt cca gaa aac ggc tga                         369
Pro Gln Gly Gly Asp Ser Pro Glu Asn Gly  *
        115                 120


<210> 52
<211> 122
<212> PRT
<213> human

<400> 52
Met Glu Val Glu Pro Ala Ala Met Val Lys Val Thr Leu Ser Pro Pro
 1               5                  10                  15
Leu Phe Ser Leu Pro Glu Phe Trp Asn Leu Ala Gly Pro Pro Ala Pro
            20                  25                  30
Val His Thr Ala Thr Gly Arg Val Thr Val Ile Gly Ile Trp Gln Asp
        35                  40                  45
Arg Arg Asp Tyr Ile Cys Pro Tyr Cys Gln Leu Lys Ala Glu Leu Cys
    50                  55                  60
Gln Gln Ile Ser Arg Glu Ala Val Glu Gln Pro Leu Ala Gly Gln Lys
65                  70                  75                  80
Ala Ala Phe Lys Thr Ala Gln Thr Ala Pro Ala Thr Gly Pro Met Thr
                85                  90                  95
Ala Ala Ala Cys Glu His Lys Ala Val Gly Ala Ala Ala Ala Phe Leu
            100                 105                 110
Pro Gln Gly Gly Asp Ser Pro Glu Asn Gly
        115                 120
```

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 April 2005 (14.04.2005)

PCT

(10) International Publication Number
WO 2005/032328 A3

(51) International Patent Classification[7]: C12Q 1/68, C07H 21/02, 21/04

(21) International Application Number: PCT/US2004/015761

(22) International Filing Date: 20 May 2004 (20.05.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/472,330 21 May 2003 (21.05.2003) US

(71) Applicant *(for all designated States except US)*: **MILLENNIUM PHARMACEUTICALS, INC.** [US/US]; 40 Landsdowne Street, Cambridge, MA 02139 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **GUILD, Braydon, C.** [US/US]; 109 Riverdale Road, Concord, MA 01742 (US). **LIAO, Hua** [CN/US]; 15 Damon Park, Apt.1, Concord, MA 02474 (US). **JONES, Michael, D.** [US/US]; 14 Schouler Court, Arlington, MA 02476 (US). **WU, Jiang** [CN/US]; 23 Judith Lane, Apartment 6, Waltham, MA 02452 (US). **ZOLG, Johannes, W.** [DE/DE]; Sportplatzweg 4A, D-82372 Weilheim (DE).

(74) Agents: **SILVERI, Jean, M.** et al.; Millennium Pharmaceuticals, Inc., 40 Landsdowne Street, Cambridge, MA 02139 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report:
15 December 2005

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF RHEUMATOID ARTHRITIS

(57) Abstract: The invention relates to compositions, kits, and methods for detecting, characterizing, preventing, and treating human Rheumatoid Arthritis (RA). A variety of newly-identified markers are provided, wherein changes in the levels of expression of one or more of the markers is correlated with RA.

WO 2005/032328 A3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US04/15761

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : C12Q 1/68; C07H 21/02, 21/04
US CL : 435/6; 536/23.1, 23.5, 24.31
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/6; 536/23.1, 23.5, 24.31

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NCBI Database, National Center for Biotechnology Information, National Library of Medicine, NIH (Bethesda, MD, USA) GenBank Accession No. NP_001054, 03 April 2003.	1-3 and 12

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
24 August 2005 (24.08.2005)

Date of mailing of the international search report
27 OCT 2005

Name and mailing address of the ISA/US
Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
Facsimile No. (703) 305-3230

Authorized officer
Carla Myers

Telephone No. 571-272-1600

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US04/15761

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-3 and 12, with respect to marker "M1"

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

BOX III. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional examination fees must be paid.

Groups 1-679, claims 1-12, drawn to methods for assessing whether a patient is afflicted with RA wherein said methods comprise determining the level of expression of one of the 679 markers set forth in Tables 1 and 2. For example, Group 1 is drawn to a method for assessing whether a patient is afflicted with RA wherein said method comprises determining the level of expression of M1 (Accession No. 4557871) as set forth in Table 1. If additional groups are elected, please specify the number of the marker to be searched.

The inventions listed as Groups 1-679 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Each of the present claims has been presented in improper Markush format, as distinct methods are improperly joined in the claims. The claims are drawn to methods for assessing whether an individual has RA by determining the level of gene expression of one of the 679 markers set forth in Tables 1 and 2. Each of the markers of Tables 1 and 2 is structurally and functionally distinct from one another and has a different special technical feature than each of the other markers. Each of the markers consists of a unique nucleotide sequence, having a different melting temperature and different specificity of hybridization. Each marker also encodes for a distinct protein having its own unique biological properties and activities. Accordingly, each marker is distinct from one another due to their diverse chemical structure, their expected different chemical properties, modes of action, and different effects. As the markers required to perform the methods of claims 1-12 do not share a special technical feature, the distinct methods may not properly be presented in the alternative. Accordingly, the claims have been separated into a number of groups corresponding to the number of different inventions encompassed by the claims, and the claims will be searched only as they read upon the invention of the elected group. Further, the methods of group I require the detection of the marker M1, accession No. 4557871. As set forth in the disclosure (Table 1 and pages 89-90), this marker, as well as the other markers set forth in Tables 1 and 2, were known in the art at the time the invention was made. As the marker of groups I-679 do not represent a contribution over the prior art, the claims lack a special technical feature that is the same as or that corresponds to a special

technical feature of the other claimed inventions. Thus, there is no special technical feature linking the recited Groups, as would be necessary to fulfill the requirement for unity of invention.

Continuation of B. FIELDS SEARCHED Item 3:
DIALOG: MEDLINE, CA, BIOSIS, EMBASE, SCISEARCH; WEST: USPT, DWPI, EPAB, JP, PGPUB
search terms: PRO1557, TF, transferring, rheumatoid arthritis, RA